

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:15:50 ; Search time 44.87 Seconds
(without alignments)
20.267 Million cell updates/sec

Title: US-09-443-986A-1
Sequence: 1 RTRLRNHHSHKANT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- A_Geneseq_0601.*
- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
 - 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
 - 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
 - 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
 - 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
 - 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
 - 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
 - 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
 - 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
 - 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
 - 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
 - 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
 - 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
 - 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
 - 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
 - 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
 - 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
 - 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
 - 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
 - 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
 - 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
 - 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	79	100.0	16 21 AAB03865	GIT receptor target
2	42	53.2	191 21 AAG36896	Arabidopsis thalia
3	42	53.2	209 21 AAG36895	Arabidopsis thalia
4	42	53.2	244 21 AAG36894	Arabidopsis thalia
5	42	53.2	618 21 AAB56803	Human prostate can
6	39	49.4	475 22 AAB64387	Amino acid sequenc
7	38	48.1	175 21 AAY70225	Human RNA-associat
8	38	48.1	175 22 AAB49669	Human ribosomal su
9	38	48.1	184 20 AAY32122	Human LYST interac
10	38	48.1	244 21 AAG07361	Arabidopsis thalia
11	38	48.1	244 21 AAG61263	Arabidopsis thalia

12	38	48.1	249 21 AAB54170	Human pancreatic c
13	38	48.1	251 21 AAG07360	Arabidopsis thalia
14	38	48.1	251 21 AAG61262	Arabidopsis thalia
15	38	48.1	258 21 AAG07359	Arabidopsis thalia
16	38	48.1	258 21 AAG61261	Arabidopsis thalia
17	38	48.1	514 19 AAW80400	A secreted protein
18	38	48.1	533 21 AAG30609	Arabidopsis thalia
19	38	48.1	624 21 AAG30608	Arabidopsis thalia
20	38	48.1	702 21 AAG30607	Arabidopsis thalia
21	38	48.1	720 22 AAY72081	Nicotiana tabacum
22	38	48.1	721 22 AAY72082	Nicotiana tabacum
23	37	46.8	305 21 AAY56330	Neisseria gonorrhoe
24	37	46.8	559 18 AAW89803	Staphylococcus aur
25	37	46.8	572 11 AAR05042	Heat shock protein
26	37	46.8	572 20 AAY23921	Amino acid sequenc
27	37	46.8	940 11 AAR07070	Fibronectin-bindin
28	37	46.8	1018 9 AAF82115	Fibronectin bindin
29	37	46.8	1027 18 AAW89806	Staphylococcus aur
30	36	45.6	68 21 AAG57415	Arabidopsis thalia
31	36	45.6	78 21 AAG57414	Arabidopsis thalia
32	36	45.6	80 21 AAG57413	Arabidopsis thalia
33	36	45.6	104 19 AAW38532	S. pneumoniae RESD
34	36	45.6	107 19 AAW57218	Desulfovibrio vulg
35	36	45.6	107 19 AAW57219	Cytochrome c3 prot
36	36	45.6	107 19 AAW57221	Cytochrome c3 prot
37	36	45.6	107 21 AAB10915	D. vulgaris cytoch
38	36	45.6	107 21 AAB10916	D. vulgaris cytoch
39	36	45.6	107 21 AAY5095	Mutant cytochrome
40	36	45.6	107 21 AAY5096	Mutant cytochrome
41	36	45.6	107 21 AAY5099	Wild type cytochro
42	36	45.6	113 21 AAG03200	Human secreted pro
43	36	45.6	133 20 AAY04842	Mycobacterium spec
44	36	45.6	135 21 AAG05473	Arabidopsis thalia
45	36	45.6	145 21 AAG05472	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
AAB03865	
ID AAB03865	standard; peptide; 16 AA.
AC AAB03865;	
XX	
XX	
DT 20-OCT-2000	(first entry)
XX	
XX	
DE	GIT receptor targeting peptide 2Elan144 (PAX2 fragment).
XX	
XX	
KW	Retros-inversion peptide; gastrointestinal tract; active agent transport;
KW	GIT; hypertension; diabetes; osteoporosis; haemophilia; anaemia; cancer;
KW	migraine; angina pectoris.
XX	
OS	Synthetic.
XX	
XX	
PH	Key
FT	Location/Qualifiers
FT	Modified-site 1
FT	/note= "Lysine is dansylated"
FT	Misc-difference 2..16
FT	/note= "D-form residue"
XX	
XX	
PN	WO200031123-A2.
XX	
PD	
XX	02-JUN-2000.
XX	
PF	19-NOV-1999; 99WO-IE00117.
XX	
XX	
PR	19-NOV-1998; 98US-0109038.
XX	
XX	(ELAN-) ELAN CORP PLC.
PA	
XX	
XX	O'Mahony DJ;
XX	

DR WPI; 2000-400037/34.
 XX Retro-inverted peptide used to deliver active agents across the
 PT gastrointestinal tract to treat hypertension, diabetes, osteoporosis,
 PT haemophilia, anaemia, cancer, migraines and angina pectoris -
 XX
 XX Disclosure; Page 18; 36pp; English.
 XX This invention relates to retro-inverted peptides which specifically bind
 CC to the gastro-intestinal tract receptor HPT1, hPPT1, D2H or hSR. Also
 CC included in the invention are a retro-inverted peptide which enhances the
 CC delivery of an active agent across the gastrointestinal tract (GIT) into
 CC the systemic, portal or hepatic circulation. A composition comprising a
 CC retro-inverted peptide bound to a material comprising an active agent
 CC used to treat a mammalian disease or disorder is also disclosed in the
 CC invention. The retro-inversion peptides target gastrointestinal tract
 CC transport receptors to promote in vivo uptake of active agents and/or
 CC enhance active agent delivery across the tract into the systemic
 CC circulation. The gastrointestinal agents (containing retro-inverted
 CC peptides) are used to facilitate the transport of active ingredients
 CC through human or animal gastrointestinal tissue, from the lumen to the
 CC portal, hepatic, or systemic circulation. The compositions containing
 CC these agents can be used to treat or prevent mammalian, especially human,
 CC diseases or disorders, especially hypertension, diabetes, osteoporosis,
 CC haemophilia, anaemia, cancer, migraine, and angina pectoris. The
 CC compositions can be administered in vivo to image selected sites or
 CC tissues, such as the gastrointestinal tract, by using an imaging agent as
 CC the active agent.
 CC The present sequence represents a retro-inversion used in the invention.
 CC The sequence is a PAX2 15 mer fragment D form retro-inversion peptide.
 XX
 SQ Sequence 16 AA;

Query Match 100.0%; Score 79; DB 21; Length 16;
 Best Local Similarity 100.0%; Pred. No. 3.7e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RTRLRRHSHSKANT 15
 DB 2 rtrlrrhshskant 16
 ||||| |||||

RESULT 2
 AAG36896
 ID AAG36896 standard; Protein; 191 AA.
 XX
 AC AAG36896;
 XX
 XX 18-OCT-2000 (first entry)
 XX
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 45280.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 XX 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999; 99US-0121825.
 XX 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR

PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 23-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.

PR	13-OCT-1999;	99US-01592995
PR	14-OCT-1999;	99US-01593290
PR	14-OCT-1999;	99US-01593300
PR	14-OCT-1999;	99US-01593301
PR	14-OCT-1999;	99US-01593637
PR	14-OCT-1999;	99US-01596378
PR	18-OCT-1999;	99US-01595884
PR	21-OCT-1999;	99US-01607441
PR	21-OCT-1999;	99US-01607677
PR	21-OCT-1999;	99US-01607688
PR	21-OCT-1999;	99US-01607700
PR	21-OCT-1999;	99US-01608174
PR	21-OCT-1999;	99US-01608175
PR	22-OCT-1999;	99US-01609801
PR	22-OCT-1999;	99US-01609802
PR	22-OCT-1999;	99US-01609809
PR	25-OCT-1999;	99US-01614404
PR	25-OCT-1999;	99US-01614405
PR	25-OCT-1999;	99US-01614406
PR	26-OCT-1999;	99US-01613590
PR	26-OCT-1999;	99US-01613629
PR	26-OCT-1999;	99US-01613661
PR	28-OCT-1999;	99US-01619920
PR	28-OCT-1999;	99US-01619992
PR	28-OCT-1999;	99US-01619993
PR	29-OCT-1999;	99US-01621412
PR	29-OCT-1999;	99US-01621422

Qy	2	TLRRNHSSHKANT	15
Dp	165	trlrngepekmnt	178

AAG36895
ID AAG36895 standard; Protein; 209 AA.

18-OCT-2000 (first entry)

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX
PN
EP1033405-A2

25-FEB-2000; 2000EP-0301439.

XX	25-FEB-1999;	9905-01218225;
	05-MAR-1999;	9905-0123180;
PR	09-MAR-1999;	9905-0123548;
PR	23-MAR-1999;	9905-0125788;
PR	25-MAR-1999;	9905-0126264;
PR	29-MAR-1999;	9905-0126785;
PR	01-APR-1999;	9905-0127462;
PR	06-APR-1999;	9905-0128234;
PR	08-APR-1999;	9905-0128714;
PR	16-APR-1999;	9905-0129845;
PR	19-APR-1999;	9905-0130077;
PR	21-APR-1999;	9905-0130449;
PR	23-APR-1999;	9905-0130510;
PR	23-APR-1999;	9905-0130891;
PR	28-APR-1999;	9905-0131449;

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140384.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142134.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

```
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161004.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.2%; Score 42; DB 21; Length 209;
Best Local Similarity 64.3%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TRLRRNHSHKANT 15
    ||||| | |
Db 183 trlrngepekmnt 196

RESULT 4
AAG36894
ID AAG36894 standard; Protein; 244 AA.
XX AC AAG36894;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 45278.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PP 06-SEP-2000.
XX PD 25-FEB-2000; 2000EP-0301439.
XX PF 25-FEB-1999; 99US-0121925.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 06-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 10-JUN-1999; 99US-0139119.
PR 14-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140595.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
```

```
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.2% Score 42; DB 21; Length 244;
Best Local Similarity 64.3%; Pred No. 12;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TRLRRNHSHKANT 15
Db 218 tlrngepekmnt 231

RESULT 5
AAB56803
ID AAB56803 standard; Protein; 618 AA.
XX AC AAB56803;
XX DT 13-MAR-2001 (first entry)
XX DE Human prostate cancer antigen protein sequence SEQ ID NO:1381.
XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.
XX OS Homo sapiens.
XX PN WO200055174-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US05988.
XX PI 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX PI Rosen CA, Ruben SM;
XX WI WPI; 2000-587513/55.
XX DR N-PSDB; AAF16006.
XX PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX Claim 11; Page 1805-1808; 2338pp; English.
XX CC AAF15566 to AAF16505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX CC The prostate cancer antigens can have neuroprotective, cytostatic,
XX CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
XX CC and can be used in gene therapy. The prostate cancer antigen
XX CC polynucleotides may be used for detection of prostate cancer, chromosome
XX CC identification, as chromosome markers, and for numerous other diagnostic
XX CC or research purposes. The prostate cancer antigens may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
```

CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.

XX Sequence 618 AA;

Query Match 53.2%; Score 42; DB 21; Length 618;
 Best Local Similarity 52.9%; Pred. No. 33;
 Matches 9; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 1 RTRLRRHSS--HKANT 15

DB 579 rrrrrrrhssggghkss 595

RESULT 6

AAAB64387
 ID AAB64387 standard; Protein; 475 AA.

XX AAB64387;

XX 22-MAR-2001 (first entry)

XX Amino acid sequence of human intracellular signalling molecule INTRA19.

XX Human; intracellular signalling molecule; INTRA; immunosuppressive;
 KW cytosolic; neuroprotective; nontropic; antiarteriosclerotic; cancer;
 KW antiinflammatory; anti-HIV; neuroleptic; antibacterial; antifungal;
 KW antiviral; antiparasitic; antihelminthic; antiparkinsonian; AIDS;
 KW cell proliferative disorder; arteriosclerosis autoimmune; epilepsy;
 KW inflammatory disorder; Addison's disease; gastrointestinal disorder;
 KW neurological disorder; Parkinson's disease; Creutzfeldt-Jakob disease;
 KW mental disorder; schizophrenia; anxiety; chromosome 7.

XX Homo sapiens.

XX WO200077040-A2.

XX 21-DEC-2000.

XX 16-JUN-2000; 2000WO-US16636.

XX 16-JUN-1999; 99US-0139566.

XX 17-AUG-1999; 99US-0149640.

XX 09-NOV-1999; 99US-0164417.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Tang YT, Hillman JL, Lal P, Bandman O, Baughn MR;

XX Azimzai Y, Yang J, Reddy R, Lu DAM;

XX WPI; 2001-025334/03.

XX N-PSDB; AAF32656.

XX New human intracellular signaling molecules, useful for the diagnosis,

XX prevention and treatment of cell proliferative, autoimmune,

XX inflammatory, neurological, gastrointestinal, reproductive and

XX developmental disorders -

XX Claim 5; Page 131-132; 192pp; English.

XX Sequences AAF32638 - AAF32689 represent cDNA encoding human

XX intracellular signalling molecules INTRA1 - INTRA52, represented in

XX AAB64369 - AAB64420. Modulators of the intracellular signalling molecules

XX of the invention exhibit immunosuppressive; cytosolic; neuroprotective;

XX nontropic; antiarteriosclerotic; antiinflammatory; anti-HIV;

XX neuroleptic; antibacterial; antifungal; antiviral; antiparasitic;

XX antihelminthic; and antiparkinsonian activity. INTRA polypeptides their

XX agonists and antagonists are useful for the treatment of a condition

XX associated with decreased or increased expression of functional INTRA.

XX Disorders associated with abnormal INTRA expression or activity include

CC cell proliferative disorders e.g. arteriosclerosis and cancers;
 CC autoimmune or inflammatory disorders e.g. Addison's disease and acquired
 CC immunodeficiency syndrome (AIDS); viral, bacterial, fungal, parasitic,
 CC protozoal and helminthic infections; gastrointestinal disorders e.g.
 CC dysphagia and irritable bowel syndrome; neurological disorders e.g.
 CC epilepsy and Parkinson's disease; prion diseases e.g. Creutzfeldt-Jakob
 CC disease and mental disorders e.g. anxiety, schizophrenia and Tourette's
 CC disorder. Antibodies immuno specific for the INTRA proteins may also be
 CC useful in the diagnosis of the above disorders.

XX Sequence 475 AA;

Query Match 49.4%; Score 39; DB 22; Length 475;

Best Local Similarity 70.0%; Pred. No. 80;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TRLRRNHSS 11

DB 57 srhrnhsqh 66

RESULT 7

AAAY70225

ID AAY70225 standard; Protein; 175 AA.

XX AAY70225;

XX 06-JUN-2000 (first entry)

XX Human RNA-associated protein-6 (RNAAP-6).

XX RNA-associated protein; RNAAP; human; clone 002690; cytostatic;
 KW immunosuppressive; antiinflammatory; keratolytic; neuroprotective;
 KW antiarteriosclerotic; hepatotropic; antipsoriatic; virucide; anti-HIV;
 KW antiallergic; antirheumatic; antiarthritic; ophthalmological; autoimmune;
 KW antimicrobial; cell proliferative disorder; inflammation; cirrhosis;
 KW actinic keratosis; bursitis; arteriosclerosis; artherosclerosis;
 KW hepatitis; myelofibrosis; primary thrombocythemia; psoriasis; cancer;
 KW mixed connective tissue disease; MCTD; HIV; uveitis; Crohn's disease;
 KW allergy; rheumatoid arthritis; parasitic infection.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Region 28..126

XX Modified-site 55 /note= "Signature sequence of ribosomal protein L17"

XX Modified-site 81 /note= "Potential phosphorylation site"

XX Modified-site 112 /note= "Potential phosphorylation site"

XX Modified-site 140 /note= "Potential glycosylation site"

XX Modified-site 164 /note= "Potential phosphorylation site"

XX Modified-site 164 /note= "Potential glycosylation site"

XX WO200011171-A2.

XX 02-MAR-2000.

XX 20-AUG-1999; 99WO-US19361.

XX 21-AUG-1998; 98US-0097550.

XX 12-JAN-1999; 99US-0115639.

XX (INCY-) INCYTE PHARM INC.

XX Hillman JL, Yue H, Tang YT, Corley NC, Guegler KJ, Gorgone GA;

XX Patterson C, Baughn MR, Lal P, Bandman O, Reddy R, Azimzai Y;

XX Shih LL, Yang J, Lu DAM;

DR WPI; 2000-237651/20.
 XX N-PSDB; AAZ51255.
 PT Human RNA-associated proteins useful in diagnosing, treating and
 PT preventing cell proliferative, autoimmune, inflammatory and infectious
 disorders -
 XX
 PS Claim 1; Page 83-84; 123pp; English.
 XX
 CC The present amino acid sequence is the human RNA-associated protein-6
 CC (RNAAP-6), identified in incyte clone 002690, derived from HMC1NOT01
 CC library. It is expressed in reproductive, nervous and cardiovascular
 CC tissues. It has cytostatic, immunosuppressive, keratolytic,
 CC antiinflammatory, antiarteriosclerotic, hepatotropic, virucide,
 CC neuroprotective, antipsoriatic, anti-HIV, antiallergic, antirheumatic,
 CC antiarthritic, ophthalmological and antimicrobial activity.
 CC RNAAP antibodies are useful for diagnosis of diseases associated with
 CC altered expression or activity of RNAAP. It is used to treat cell
 CC proliferative, autoimmune, inflammatory and infectious disorders, like
 CC actinic keratosis, bursitis, arteriosclerosis, artherosclerosis,
 CC cirrhosis, hepatitis, myelofibrosis, mixed connective tissue disease
 CC (MCTD), psoriasis, primary thrombocythemia and cancer, HIV, allergies,
 CC rheumatoid arthritis, uveitis, Crohn's disease, and bacterial, viral and
 CC parasitic infections.
 XX
 SQ Sequence 175 AA;
 Query Match 48.1%; Score 38; DB 21; Length 175;
 Best Local Similarity 77.8%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 7 NHSSHKANT 15
 Db 164 hnshtaqt 172
 RESULT 8
 AAB49669
 ID AAB49669 standard; Protein; 175 AA.
 AC AAB49669;
 XX
 XX 27-MAR-2001 (first entry)
 DT Human ribosomal subunit protein h60srPL17 amino acid sequence.
 DE
 XX Human; ribosomal subunit protein; h60srPL17; adrenal gland.
 KW
 XX Homo sapiens.
 OS
 PN CNI268570-A.
 XX
 PD 04-OCT-2000.
 XX
 PF 02-MAR-2000; 2000CN-0111776.
 XX
 XX 02-MAR-2000; 2000CN-0111776.
 PR
 XX (SCHR-) SOUTH CHINA RES CENT NAT HUMAN GENE GROU.
 PA
 XX Li N, Qian B, Peng Y;
 PI
 XX WPI; 2001-050474/07.
 DR
 XX N-PSDB; AAC92511.
 XX
 PT New human ribosomal subunit protein and its coding sequence -
 PS
 PS Claim 2; Fig 2; 20pp; Chinese.
 XX
 CC This invention relates to a human ribosomal subunit protein h60srPL17
 CC represented by sequence AAB49669, which is encoded by a cDNA sequence
 CC represented in AAC92511. The ribosomal subunit protein is expressed in

CC the adrenal gland. The invention includes methods for the preparation of
 CC h60srPL17 protein and nucleic acid sequences, and a method for detecting
 CC h60srPL17 in a sample.
 XX
 SQ Sequence 175 AA;
 Query Match 48.1%; Score 38; DB 22; Length 175;
 Best Local Similarity 77.8%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 7 NHSSHKANT 15
 Db 164 hnshtaqt 172
 RESULT 9
 AAY32122
 ID AAY32122 standard; Protein; 184 AA.
 XX
 AC AAY32122;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human LYST interacting protein LIP2.
 XX
 KW LIP2; human; LYST-2; LYST interacting protein; lysosome; CHS;
 KW Chediak-Higashi syndrome; autoimmune disease; rheumatoid arthritis;
 KW systemic lupus erythematosus; inflammatory bowel disease;
 KW diabetes mellitus; multiple sclerosis; atopic disease; asthma;
 KW hay fever; rhinitis; urticaria; nasal polyp; cancer;
 KW neurodegenerative disease; pigmentation disorder; viral disease;
 KW platelet dysfunction.
 XX
 OS Homo sapiens.
 XX
 PN WO9951741-A2.
 XX
 PD 14-OCT-1999.
 XX
 PF 29-MAR-1999; 99WO-US06831.
 XX
 PR 03-APR-1998; 98US-0054956.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Nandabalan K, Kingsmore S;
 XX
 DR WPI; 1999-620203/53.
 DR N-PSDB; AAZ34488.
 XX
 PT Protein complexes, interacting proteins, and related polynucleotides
 PT useful for treating and preventing e.g. atopic, autoimmune or
 PT neurodegenerative diseases -
 XX
 PS Claim 20; Fig 3; 172pp; English.
 XX
 CC The present sequence represents a novel human LYST interacting
 CC protein, LIP2, that shows similarity to the ribosomal protein L17.
 CC LYST is the human lysosomal Chediak-Higashi syndrome (CHS) protein.
 CC The invention relates to complexes of LYST or LYST-2 (see AAY32120)
 CC with proteins identified as interacting with LYST or LYST-2 by a
 CC modified yeast two-hybrid assay system. The interacting proteins
 CC include 10 novel proteins, LIP1-10 (see AAY32121-30). Methods of
 CC screening the protein complexes for efficacy in treating and/or
 CC preventing atopic diseases (e.g. asthma, nasal polyps, hay fever
 CC rhinitis, urticaria) autoimmune diseases (e.g. CHS, rheumatoid
 CC arthritis, systemic lupus erythematosus, inflammatory bowel disease,
 CC diabetes mellitus, multiple sclerosis), neurodegenerative disease,
 CC certain cancers, pigmentation disorders, platelet dysfunction and
 CC viral diseases are provided. Nucleic acids (see AAZ34487-96)
 CC encoding LIP1-10, modulation of LIP function by gene therapy, use
 CC of antisense oligonucleotides for suppression of LIP protein

CC expression, screening for agonists and antagonists, diagnosing or
CC screening for the presence of a predisposition to a disease or
CC disorder, and animal models are also disclosed.

XX
SQ Sequence 184 AA;

Query Match 48.1%; Score 38; DB 20; Length 184;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 NHSSHKANT 15
|||||
Db 173 nhshtaq 181

RESULT 10

AAG07361
ID AAG07361 standard; Protein: 244 AA.

XX
AC AAG07361;

XX
DT 17-OCT-2000 (first entry)

XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4485.

XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX
OS Arabidopsis thaliana.

XX
PN EPI033405-A2.

XX
PD 06-SEP-2000.

XX
PF 25-FEB-2000; 2000EP-0301439.

XX
PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

```
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 48.1%; Score 38; DB 21; Length 244;
Best Local Similarity 54.5%; Pred. No. 58;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LRRNHSSHKAN 14
Db 7 lsrnhqshrss 17
| ||| ||:::
RESULT 11
AAG61263
ID AAG61263 standard; Protein; 244 AA.
XX
AC AAG61263;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 79443.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
```

PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 48.1%; Score 38; DB 21; Length 244;
Best Local Similarity 54.5%; Pred. No. 58;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 4 LRRNHSSKAN 14
| | | | | | | | | |
Db 7 lsrnqshrss 17

RESULT 12
AAB54170
ID AAB54170 standard; Protein; 249 AA.
XX AAB54170;
XX
XX
XX 09-MAR-2001 (first entry)
XX Human pancreatic cancer antigen protein sequence SEQ ID NO:622.
XX
XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW detection; diagnosis; identification; cytostatic; neuroprotective;
KW neoplastic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic;
KW neural; immune system; muscular; reproductive; gastrointestinal;
KW pulmonary; cardiovascular; renal; proliferative.
XX
XX Homo sapiens.
XX
XX WO200055320-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05989.
XX
XX 12-MAR-1999; 99US-0124270.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
XX WPI; 2000-579444/54.
XX N-PSDB; AAC98935.
XX
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition -
XX
XX Claim 11; Page 1060-1061; 1379pp; English.
XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, neoplastic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
XX Sequence 249 AA;

Query Match 48.1%; Score 38; DB 21; Length 249;
Best Local Similarity 63.6%; Pred. No. 60;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 RRHSHSKANT 15
||| :||:|
Db 236 rrnptshragt 246
RESULT 13
AAG07360
ID AAG07360 standard; Protein; 251 AA.
XX AAG07360;
XX
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 4484.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 48.1%; Score 38; DB 21; Length 251;
Best Local Similarity 54.5%; Pred. No. 60;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 LRRNHSHKAN 14
| | | | | | | | | |
Db 14 lsrnhqshrs 24

RESULT 14
AAG61262

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145216.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 48.1%; Score 38; DB 21; Length 258;
Best Local Similarity 54.5%; Pred. No. 62;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LRRNHSSHKAN 14
| | | | |
Db 21 lsruhghsrss 31

Search completed: November 21, 2001, 16:20:49
Job time: 299 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:16:11 ; Search time 24.83 Seconds
(without alignments)
13.594 Million cell updates/sec

Title: US-09-443-986A-1

Perfect score: 79

Sequence: 1 RTRLRHHSHKANT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	48.1	514	2	US-08-960-022-14
2	37	46.8	45	2	US-08-472-244-7
3	37	46.8	52	1	US-08-294-189-17
4	36	45.6	208	4	US-09-094-103-4
5	36	45.6	232	4	US-09-094-103-2
6	36	45.6	235	2	US-08-411-607A-2
7	36	45.6	308	3	US-08-927-433-9
8	36	45.6	308	4	US-08-775-882-4
9	36	45.6	308	4	US-08-775-882-6
10	36	45.6	424	2	US-08-614-155B-3
11	36	45.6	570	4	US-08-991-944-4
12	35	44.3	520	4	US-08-810-009-5
13	35	44.3	601	1	US-08-458-477A-2
14	35	44.3	601	2	US-09-033-153-2
15	35	44.3	601	4	US-09-325-430B-2
16	35	44.3	1253	1	US-08-252-966B-12
17	35	44.3	1251	1	US-08-252-966B-18
18	35	44.3	2332	1	US-07-864-004B-4
19	35	44.3	2332	1	US-08-251-937A-4
20	35	44.3	2332	1	US-08-212-133A-2
21	35	44.3	2332	1	US-08-276-594A-2
22	35	44.3	2332	1	US-08-474-503-2
23	35	44.3	2332	2	US-08-670-707A-2
24	35	44.3	2332	4	US-09-037-601-2
25	35	44.3	2332	4	US-09-324-867-3
26	35	44.3	2332	5	PCT-US93-03275-4
27	35	44.3	2332	5	PCT-US94-13200-2

28	35	44.3	2351	1	US-08-121-202-2
29	35	44.3	2351	1	US-08-366-851A-2
30	35	44.3	2351	6	5171844-2
31	35	44.3	2351	6	5422260-1
32	34	43.0	47	4	US-08-905-223-350
33	34	43.0	370	4	US-09-199-737-2
34	34	43.0	416	1	US-08-252-995D-2
35	34	43.0	416	2	US-08-834-108-2
36	34	43.0	451	2	US-09-016-000-3
37	34	43.0	464	1	US-08-252-995D-6
38	34	43.0	464	2	US-08-834-108-6
39	34	43.0	831	2	US-09-047-026A-4
40	34	43.0	925	1	US-08-252-995D-4
41	34	43.0	925	2	US-08-834-108-4
42	34	43.0	1311	2	US-08-540-406-4
43	34	43.0	1311	3	US-08-656-055-4
44	34	43.0	1311	4	US-08-854-668-4
45	34	43.0	1311	5	PCT-US95-13233-4

ALIGNMENTS

RESULT 1
US-08-960-022-14
; Sequence 14, Application US/08960022
; Patent No. 5976837
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,022
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 514 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-960-022-14

Query Match 48.1%; Score 38; DB 2; Length 514;
Best Local Similarity 54.5%; Pred. No. 82;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LRRNHSHKA 13
I:|:| I I
Db 39 RIRSHQKHA 49

RESULT 2

US-08-472-244-7
; Sequence 7, Application US/08472244
; Patent No. 5821088
; GENERAL INFORMATION:
; APPLICANT: Darzins, Aldis
; APPLICANT: Whitehead, Stephen
; APPLICANT: Hruba, Dennis E.
; TITLE OF INVENTION: Use of Gram-Positive Bacteria to Express
; TITLE OF INVENTION: Recombinant Proteins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314-3187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,244
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 016921-065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; STRAIN: staphylococcal fibronectin-binding protein

US-08-472-244-7

Query Match 46.8%; Score 37; DB 2; Length 45;
Best Local Similarity 70.0%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LRRNHSHKA 13
I:|:| I:|:|
Db 36 LRRNKNHKA 45

RESULT 3

US-08-294-189-17
; Sequence 17, Application US/08294189
; Patent No. 5707822
; GENERAL INFORMATION:
; APPLICANT: Fischetti, Vincent A.
; APPLICANT: Rakonjac, Jasna
; APPLICANT: Robbins, John
; TITLE OF INVENTION: GENE SERUM OPACITY FACTOR
; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/294,189
; FILING DATE: 22-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-118
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus aureus

US-08-294-189-17

Query Match 46.8%; Score 37; DB 1; Length 52;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LRRNHSHKA 13
I:|:| I:|:|
Db 43 LRRNKNHKA 52

RESULT 4

US-09-094-103-4
; Sequence 4, Application US/09094103
; Patent No. 6140061
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; APPLICANT: Zalacain, Magdalena
; APPLICANT: Throup, John
; APPLICANT: Biswas, Sanjoy
; TITLE OF INVENTION: Response regulator
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,103
; FILING DATE:

```
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,332
; FILING DATE: 20-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10016
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-094-103-4

Query Match 45.6%; Score 36; DB 4; Length 208;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 LRRNHSHK 12
Db 111 LKNHSSHE 119

RESULT 5
US-09-094-103-2
; Sequence 2, Application US/09094103
; Patent No. 6140061
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; APPLICANT: Zalacain, Magdalena
; APPLICANT: Throup, John
; APPLICANT: Biswas, Sanjoy
; TITLE OF INVENTION: Response regulator
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,103
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,332
; FILING DATE: 20-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10016
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; US-09-094-103-2

Query Match 45.6%; Score 36; DB 4; Length 232;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 LRRNHSHK 12
Db 135 LKNHSSHE 143

RESULT 6
US-08-411-607A-2
; Sequence 2, Application US/08411607A
; Patent No. 5994102
; GENERAL INFORMATION:
; APPLICANT: HUDSON, PETER L
; APPLICANT: ROSEN, CRAIG A
; APPLICANT: HE, WEI WU
; TITLE OF INVENTION: PROSTATIC GROWTH FACTOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NJ
; COUNTRY: US
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/411,607A
; FILING DATE: 11-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-329
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)994-1700
; TELEFAX: (201) 994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-411-607A-2

Query Match 45.6%; Score 36; DB 2; Length 295;
Best Local Similarity 57.1%; Pred. No. 99;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 TRLRRNHSHKANT 15
Db 52 TRLRANQWEDSNT 65

RESULT 7
US-08-927-433-9
; Sequence 9, Application US/08927433
; Patent No. 6107476
; GENERAL INFORMATION:
; APPLICANT: Erlander, Mark G.
; APPLICANT: Huang, Shaoming
; APPLICANT: Jackson, Michael A.
```

APPLICANT: Peterson, Per A.
TITLE OF INVENTION: PROTEIN GROWTH FACTOR FOR TREATING PROSTATE CANCER, AND RELATED
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Johnson & Johnson
STREET: One J & J Plaza
CITY: New Brunswick
STATE: New Jersey
COUNTRY: USA
ZIP: 08933
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,433
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Morrison, Alan J.
REGISTRATION NUMBER: 37,399
REFERENCE/DOCKET NUMBER: ORT-849
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-524-3592
TELEFAX: 732-524-2808
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-927-433-9

Query Match 45.6%; Score 36; DB 3; Length 308;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TRLRRNHSHKANT 15
||||| :||
Db 65 TRLRANQSWEDSNT 78

RESULT 8
US-08-775-882-4
Sequence 4, Application US/08775882
Patent No. 6180602
GENERAL INFORMATION:
APPLICANT: KATO, Seishi
APPLICANT: OH, Suwan
APPLICANT: SEKINE, Shingo
APPLICANT: SAEKI, Mihiro
APPLICANT: KOBAYASHI, Midori
APPLICANT: YADA, Mika
APPLICANT: TSUJI, Tomoko
APPLICANT: OHMORI, Hitoshi
TITLE OF INVENTION: HUMAN NOVEL CDNA, TGF-BETA SUPERFAMILY
TITLE OF INVENTION: PROTEIN ENCODED THEREBY AND THE USE AS IMMUNOSUPPRESSIVE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,882
FILING DATE: 02-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,207
FILING DATE: 16-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/379,441
FILING DATE: 03-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-61431
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-327619
FILING DATE: 13-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-208077
FILING DATE: 04-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: WEGNER, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 53466/194
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-882-4

Query Match 45.6%; Score 36; DB 4; Length 308;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TRLRRNHSHKANT 15
||||| :||
Db 65 TRLRANQSWEDSNT 78

RESULT 9
US-08-775-882-6
Sequence 6, Application US/08775882
Patent No. 6180602
GENERAL INFORMATION:
APPLICANT: KATO, Seishi
APPLICANT: OH, Suwan
APPLICANT: SEKINE, Shingo
APPLICANT: SAEKI, Mihiro
APPLICANT: KOBAYASHI, Midori
APPLICANT: YADA, Mika
APPLICANT: TSUJI, Tomoko
APPLICANT: OHMORI, Hitoshi
TITLE OF INVENTION: HUMAN NOVEL CDNA, TGF-BETA SUPERFAMILY
TITLE OF INVENTION: PROTEIN ENCODED THEREBY AND THE USE AS IMMUNOSUPPRESSIVE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,882
FILING DATE: 02-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,207
FILING DATE: 16-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/379,441
FILING DATE: 03-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-61431
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-327619
FILING DATE: 13-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-208077
FILING DATE: 04-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: WEGNER, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 53466/194
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-882-6

Query Match 45.6%; Score 36; DB 4; Length 308;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TRLRRNHSHKANT 15
||||| :||
Db 65 TRLRANQSWEDSNT 78

RESULT 10
US-08-614-155B-3
Sequence 3, Application US/08614155B
Patent No. 5919692
GENERAL INFORMATION:
APPLICANT: Susumu SEINO et al.
TITLE OF INVENTION: NOVEL UBIQUITOUS POTASSIUM-CHANNEL
TITLE OF INVENTION: PROTEINS AND THEIR GENES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC
STREET: 400 Seventh Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/614,155B
FILING DATE: 12-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 264942/1995
FILING DATE: 18-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: PLAYER, William E.
REGISTRATION NUMBER: 31,409
REFERENCE/DOCKET NUMBER: 10890/P60281US0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-638-6666
TELEFAX: 202-393-5350
TELEX: RCA 248593 IDEA UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-614-155B-3

Query Match 45.6%; Score 36; DB 2; Length 424;
Best Local Similarity 40.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RTRLRRNHSHKANT 15
| :|||:| :||
Db 383 RNSMRNNSMRRSNS 397

RESULT 11
US-08-991-944-4
Sequence 4, Application US/08991944
Patent No. 6280955
GENERAL INFORMATION:
APPLICANT: Cao, Zhaodan
TITLE OF INVENTION: Interleukin-1 Receptor Accessory Proteins,
TITLE OF INVENTION: Nucleic Acids and Methods
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,944
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-014
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 570 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-944-4

Query Match 45.6%; Score 36; DB 4; Length 570;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 4 LRNRHSHKANT 15
| | | | : | | |
Db 326 LRNVVCHARNT 337

RESULT 12
US-08-810-009-5
; Sequence 5, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SEUTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810.009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-810-009-5

Query Match 44.3%; Score 35; DB 4; Length 520;
Best Local Similarity 53.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 LRNRHSHKANT 15
| | | | : | | |
Db 44 RIRRRSGSKNT 56

RESULT 13
US-08-458-477A-2
; Sequence 2, Application US/08458477A
; Patent No. 5723311
; GENERAL INFORMATION:
; APPLICANT: WEI, ET AL.
; TITLE OF INVENTION: Human Topoisomerase I
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY

; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458.477A
; FILING DATE: June 2, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05701
; FILING DATE: 18 MAY 94
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-309 (PF118P1)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 601 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-458-477A-2

Query Match 44.3%; Score 35; DB 1; Length 601;
Best Local Similarity 53.8%; Pred. No. 2.9e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RTRLRRHSHKA 13
| | | | : | | |
Db 498 RAELRRARAHEKA 510

RESULT 14
US-09-033-153-2
; Sequence 2, Application US/09033153
; Patent No. 5968803
; GENERAL INFORMATION:
; APPLICANT: WEI, ET AL.
; TITLE OF INVENTION: Human Topoisomerase I
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/033.153
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/458,477
; FILING DATE: June 2, 1995
; APPLICATION NUMBER: PCT/US94/05701
; FILING DATE: 18 MAY 94
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-309 (PF118P1)
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 601 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-09-033-153-2

Query Match 44.3%; Score 35; DB 2; Length 601;
Best Local Similarity 53.8%; Pred. No. 2.9e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 RTRLRRNHSSHKA 13
| | | | : | | |
Db 498 RAE LRRRAAEHKA 510

RESULT 15
US-09-325-430B-2
Sequence 2, Application US/09325430B
Patent No. 6255077
GENERAL INFORMATION:
APPLICANT: Wei et al.
TITLE OF INVENTION: Human DNA Topoisomerase 1 Alpha
FILE REFERENCE: PF118D2
CURRENT APPLICATION NUMBER: US/09/325,430B
CURRENT FILING DATE: 1999-06-04
PRIOR APPLICATION NUMBER: 09/033,153
PRIOR FILING DATE: 1998-03-02
PRIOR APPLICATION NUMBER: 08/458,477
PRIOR FILING DATE: 1995-06-02
PRIOR APPLICATION NUMBER: PCT/US94/05701
PRIOR FILING DATE: 1994-05-18
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 601
TYPE: PRT
ORGANISM: Homo sapiens
US-09-325-430B-2

Query Match 44.3%; Score 35; DB 4; Length 601;
Best Local Similarity 53.8%; Pred. No. 2.9e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 RTRLRRNHSSHKA 13
| | | | : | | |
Db 498 RAE LRRRAAEHKA 510

Search completed: November 21, 2001, 16:21:20
Job time: 309 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:16:46 ; Search time 26.99 seconds
(without alignments)
42.335 Million cell updates/sec

Title: US-09-443-986A-1

Perfect score: 79

Sequence: 1 RTRLRRNHSSHKANT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	51.9	366	2 H96537	hypothetical prote
2	40	50.6	419	2 T39920	probable glucanase
3	40	50.6	1095	2 T00329	hypothetical prote
4	40	50.6	1460	2 T00095	hypothetical prote
5	39	49.4	381	2 C71680	hypothetical prote
6	39	49.4	424	2 T21181	hypothetical prote
7	39	49.4	926	2 A54142	nucleoporin NUP107
8	39	49.4	1212	2 S27771	RNA-directed DNA p
9	38	48.1	63	2 B70237	hypothetical prote
10	38	48.1	220	2 S20064	ribosomal protein
11	38	48.1	318	2 D71501	hypothetical prote
12	38	48.1	401	2 T14429	arginine decarboxy
13	38	48.1	404	2 T25420	hypothetical prote
14	38	48.1	407	2 S27774	transforming prote
15	38	48.1	408	2 T08104	arginine decarboxy
16	38	48.1	502	1 JQ2341	arginine decarboxy
17	38	48.1	605	2 S67815	protein-tyrosine k
18	38	48.1	621	2 T15046	arginine decarboxy
19	38	48.1	702	2 S71239	arginine decarboxy
20	38	48.1	702	2 A84341	arginine decarboxy
21	38	48.1	711	2 T05291	arginine decarboxy
22	38	48.1	725	2 T10721	arginine decarboxy
23	38	48.1	725	2 T10709	arginine decarboxy
24	38	48.1	1033	2 I48775	Smcx protein (esca
25	38	48.1	1043	2 A56037	DNA-binding protei
26	38	48.1	1117	2 JC4934	delta-crystallin/E
27	38	48.1	1124	2 JX0293	zinc finger protei
28	38	48.1	1154	2 A56242	E-box-binding repr
29	38	48.1	1560	2 I54361	SMCX protein - hum

30	38	48.1	7576	2 T17428	FK506 polyketide s
31	37	46.8	109	2 C72588	hypothetical prote
32	37	46.8	275	2 B48600	RNA-binding protei
33	37	46.8	419	2 F96704	hypothetical prote
34	37	46.8	439	2 A84431	probable C2H2-type
35	37	46.8	444	2 T38760	hypothetical prote
36	37	46.8	572	2 JQ0157	heat shock protein
37	37	46.8	940	2 S19702	fibronectin-bindin
38	37	46.8	1018	2 A32192	fibronectin-bindin
39	36.5	46.2	1159	2 T43461	probable phosphodi
40	36.5	46.2	1687	2 T30244	phosphodiesterase
41	36.5	46.2	1706	2 T30175	exoribonuclease, v
42	36.5	46.2	1719	2 T30174	exoribonuclease, v
43	36	45.6	1130	1 CCDV3M	cytochrome c3 prec
44	36	45.6	1134	1 R5NT16	ribosomal protein
45	36	45.6	202	2 S07979	virf protein - Agr

ALIGNMENTS

RESULT 1

H96537

hypothetical protein F14I3.21 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: H96537

R:heologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: H96537

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-366 <STO>

A:Cross-references: GB:AE005173; NID:g5734783; PIDN:AAD50048.1; GSPDB:GN00141

C:Genetics:

A:Gene: F14I3.21

A:Map position: 1

Query Match 51.9%; Score 41; DB 2; Length 366;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 LRRNHSHKA 13
|||||
Db 47 LRRNHSHGA 56

RESULT 2

T39920

probable glucanase precursor - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T39920

R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.

submitted to the EMBL Data Library, October 1997

A:Reference number: 221891

A:Accession: T39920

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-419 <MCD>

A:Cross-references: EMBL:AL121794; PIDN:CAB57923.1; GSPDB:GN00067

A:Experimental source: strain 972h-; cosmid c21B10

C:Genetics:

A:Gene: SPAC21B10.07
A:Map position: 2

Query Match 50.6%; Score 40; DB 2; Length 419;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RTRLRHHSHKANT 15
||:| | | | | | | |
Db 114 RTAVRDRHPSYKAT 128

RESULT 3

hypothetical protein KIAA0553 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00329
R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
DNA Res. 5, 31-39, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The complete
A:Reference number: 214086; MUID:98290545
A:Accession: T00329
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1095 <NAG>
A:Cross-references: EMBL:AB011125; NID:g3043629; PIDN:BAA25479.1; PID:g3043630
A:Experimental source: brain
C:Genetics:
A:Note: KIAA0553

Query Match 50.6%; Score 40; DB 2; Length 1095;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RTRLRHHSHKAN 14
|:| | | | | | | |
Db 446 RSRGRRHSHRSS 459

RESULT 4

hypothetical protein KIAA0470 - human
C:Species: Homo sapiens (man)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 01-Dec-2000
C:Accession: T00095; T47155
R:Seki, N.; Ohira, M.; Nagase, T.; Ishikawa, K.; Miyajima, N.; Nakajima, D.; Nomura, N.;
DNA Res. 4, 345-349, 1997
A:Title: Characterization of cDNA clones in size-fractionated cDNA libraries from human
A:Reference number: 214085; MUID:98116662
A:Accession: T00095
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1460 <SEK>
A:Cross-references: EMBL:AB007939; NID:g3413901; PIDN:BAA32315.1; PID:g3413902
R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: 724374
A:Accession: T47155
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1288-1460 <AAA>
A:Cross-references: EMBL:AL161964
A:Experimental source: fetal brain; clone DKFZp564B0982
C:Genetics:
A:Map position: 1
A:Note: KIAA0470; DKFZp564B0982.1

Query Match 50.6%; Score 40; DB 2; Length 1460;

Best Local Similarity 77.8%; Pred. No. 94;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 LRRNHSHK 12
| | | | | | | |
Db 400 LRRHSHK 408

RESULT 5

hypothetical protein RP255 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: C71680
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499
A:Accession: C71680
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-381 <AND>
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA14717.1; PID:g386
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: RP255

Query Match 49.4%; Score 39; DB 2; Length 381;
Best Local Similarity 63.6%; Pred. No. 38;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 LRRNHSHKAN 14
| | | | | | | |
Db 177 LHYNHTSHKLN 187

RESULT 6

hypothetical protein F21A3.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T21181
R:McMurray, A.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19387
A:Accession: T21181
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-424 <WIL>
A:Cross-references: EMBL:Z81509; PIDN:CAB04153.1; GSPDB:GN00023; CESP:F21A3.2
A:Experimental source: clone F21A3
C:Genetics:
A:Gene: CESP:F21A3.2
A:Map position: 5
A:Introns: 58/1; 119/1; 181/2; 229/1; 248/3; 314/3; 360/3
C:Superfamily: kidney bean purple acid phosphatase; phosphoesterase core homology

Query Match 49.4%; Score 39; DB 2; Length 424;
Best Local Similarity 63.6%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RTRLRHHSSH 11
| | | | | | | |
Db 409 KTRRRKHSRH 419

RESULT 7

nucleoporin NUP107 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 05-Nov-1999

C:Accession: A54142
 R:Radu, A.; Blobel, G.; Wozniak, R.W.
 J. Biol. Chem. 269, 17600-17605, 1994
 A:Title: Nup107 is a novel nuclear pore complex protein that contains a leucine zipper.
 A:Reference number: A54142; MUID:94292521
 A:Accession: A54142

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-926 <RAD>
 A:Cross-references: GB:L31840; NID:g950094; PIDN:AAA74476.1; PID:g510717

Query Match 49.4%; Score 39; DB 2; Length 926;
 Best Local Similarity 63.6%; Pred. No. 90;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 TRLRRNHSSHK 12
 || | ||| ||
 Db 20 TRTARKHSAHK 30

RESULT 8
 S27771
 RNA-directed DNA polymerase (EC 2.7.7.49) - African malaria mosquito transposon RT1 (fra
 N:Alternate names: reverse transcriptase
 C:Species: Anopheles gambiae (African malaria mosquito)
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Sep-1997
 C:Accession: S27771
 R:Besansky, N.J.; Paskewitz, S.M.; Mills-Hamm, D.M.; Collins, F.H.
 submitted to the EMBL Data Library, June 1992
 A:Description: Distinct families of site-specific retrotransposons occupy identical positions
 A:Reference number: S27770
 A:Accession: S27771
 A:Molecule type: DNA
 A:Residues: 1-1212 <BES>
 A:Cross-references: EMBL:M93690; NID:gi159615; PID:gi159617
 C:Keywords: nucleotidyltransferase

Query Match 49.4%; Score 39; DB 2; Length 1212;
 Best Local Similarity 61.5%; Pred. No. 1.2e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TRLRRNHSSHKAN 14
 ||| |||| | :
 Db 829 TRLRNHSGPKTS 841

RESULT 9
 B70237
 hypothetical protein BBH17 - Lyme disease spirochete plasmid H/lp28-3
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
 C:Accession: B70237
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
 son, D.; Peterson, J.; Kexlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
 A:Reference number: A70100; MUID:98065943
 A:Accession: B70237
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-63 <KLE>
 A:Cross-references: GB:AE000784; NID:g2690041; PIDN:AAC56021.1; PID:g2690077; TIGR:BBH17
 A:Experimental source: strain B31
 C:Genetics:
 A:Genome: plasmid

Query Match 48.1%; Score 38; DB 2; Length 63;
 Best Local Similarity 53.8%; Pred. No. 10;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 2 TRLRRNHSSHKAN 14
 : : |||| |||
 Db 13 SKIDHNHSSYKAN 25

RESULT 10
 S20064
 ribosomal protein S8.e, cytosolic - Leishmania major
 C:Species: Leishmania major
 C:Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 13-Aug-1999
 C:Accession: S20063; S20064; S24182
 R:Bard, E.; Janeczko, R.
 Nucleic Acids Res. 20, 369, 1992
 A:Title: Sequence and genomic locus of a Leishmania protein homologous to rat ribosom
 A:Reference number: S20063; MUID:92158630
 A:Accession: S20063
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-220 <BAR>
 A:Cross-references: EMBL:X62942; NID:g9577; PIDN:CAA44714.1; PID:g9578
 A:Experimental source: strain LT252
 A:Genetics: UPG
 A:Note: upstream gene
 A:Note: this is a revision to the sequence from reference S24182
 A:Accession: S20064
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-220 <BAW>
 A:Cross-references: EMBL:X62942; NID:g9577; PIDN:CAA44715.1; PID:g9579
 A:Experimental source: strain LT252
 A:Genetics: DWG
 A:Note: downstream gene
 R:Bard, E.
 Biochem. Cell Biol. 67, 516-524, 1989
 A:Title: Molecular biology of Leishmania.
 A:Reference number: S24182; MUID:90027691
 A:Accession: S24182
 A:Molecule type: DNA
 A:Residues: 1-2, M', 4-39, P', 41-43, T', 45-52, 'LRSVPCAWTPATLP', 'GLRG', 72-83, 'TPRRTSMCV
 A:Cross-references: EMBL:X62942
 A:Genetics: DWG
 A:Note: downstream gene
 A:Note: this sequence has been revised in reference S20063
 C:Genetics: <UPG>
 A:Note: upstream gene
 C:Genetics: <DWG>
 A:Note: downstream gene
 C:Superfamily: rat ribosomal protein S8
 C:Keywords: protein biosynthesis; ribosome

Query Match 48.1%; Score 38; DB 2; Length 220;
 Best Local Similarity 66.7%; Pred. No. 34;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TRLRRNHSSHKA 13
 || |||| |
 Db 163 TRRRRNHRVEKA 174

RESULT 11
 D71501
 hypothetical protein CF547 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
 C:Species: Chlamydia trachomatis
 C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 28-Jul-2000
 C:Accession: D71501
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
 Science 282, 754-759, 1998
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
 A:Reference number: A71570; MUID:99000809
 A:Accession: D71501

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-318 <ARN>
A:Cross-references: GB:AF001325; GB:AF001273; NID:g3328980; PIDN:AAC68149.1; PID:g332898
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: CT547
C:Superfamily: Chlamydia trachomatis hypothetical protein CT547

Query Match 48.1%; Score 38; DB 2; Length 318;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 4 LRRNHSHKANT 15
Db 242 LRKQHPNPSNT 253
||:|:|:|:|

RESULT 12
T14429
arginine decarboxylase (EC 4.1.1.19) - wild cabbage (fragment)
C:Species: Brassica oleracea (wild cabbage)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 03-Nov-2000
C:Accession: T14429
R:Galloway, G.L.; Malmberg, R.L.; Price, R.A.
Mol. Biol. Evol. 15, 1312-1320, 1998
A:Title: Phylogenetic utility of the nuclear gene arginine decarboxylase: an example from
A:Reference number: Z16357; MUID:99003705
A:Accession: T14429
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-401 <GAL>
A:Cross-references: EMBL:AF045683; NID:g3421189; PIDN:AAC68528.1; PID:g3421190
C:Genetics:
A:Note: spe2
C:Function:
A:Description: catalyzes decarboxylation of arginine into agmatine
A:Pathway: putrescine biosynthesis
C:Superfamily: arginine decarboxylase
C:Keywords: carbon-carbon lyase; carboxy-lyase; pyridoxal phosphate

Query Match 48.1%; Score 38; DB 2; Length 401;
Best Local Similarity 46.7%; Pred. No. 60;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RTRLRRNHSHKANT 15
Db 107 RAKLRTKHSFGST 121
|:|:|:|:|

RESULT 13
T25420
hypothetical protein T28F3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000
C:Accession: T25420
R:Mortimore, B.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z20032
A:Accession: T25420
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-404 <WIL>
A:Cross-references: EMBL:Z82285; PIDN:CAE05297.1; GSPDB:GN00022; CESP:T28F3.3
A:Experimental source: clone T28F3
C:Genetics:
A:Gene: CESP:T28F3.3
A:Map position: 4
A:Introns: 108/3; 231/3; 297/3; 339/1; 388/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C30H6.2

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-318 <ARN>
A:Cross-references: GB:AF001325; GB:AF001273; NID:g3328980; PIDN:AAC68149.1; PID:g332898
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: CT547
C:Superfamily: Chlamydia trachomatis hypothetical protein CT547

Query Match 48.1%; Score 38; DB 2; Length 404;
Best Local Similarity 55.6%; Pred. No. 60;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LRRNHSHK 12
Db 63 IRKNHTSHR 71
||:|:|:|:|

RESULT 14
S27774
transforming protein (myc) - starfish (Asterias vulgaris)
C:Species: Asterias vulgaris
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999
C:Accession: S27774
R:Walker, C.W.; Boom, J.; Marsh, A.
submitted to the EMBL Data Library, December 1991
A:Description: Novel C-myc gene is expressed seasonal in an invertebrate testis.
A:Reference number: S27774
A:Accession: S27774
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-407 <WAL>
A:Cross-references: EMBL:M80364; NID:g155850; PID:g155851
C:Superfamily: myc transforming protein; myc transforming protein homology
F:12-404/Domain: myc transforming protein homology <MYC>

Query Match 48.1%; Score 38; DB 2; Length 407;
Best Local Similarity 54.5%; Pred. No. 60;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 5 RRNHSHKANT 15
Db 210 KKHSHVHKINT 220
::|:|:|:|

RESULT 15
T08104
arginine decarboxylase (EC 4.1.1.19) - papaya (fragment)
C:Species: Carica papaya (papaya)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 03-Nov-2000
C:Accession: T08104
R:Galloway, G.L.; Malmberg, R.L.; Price, R.A.
Mol. Biol. Evol. 15, 1312-1320, 1998
A:Title: Phylogenetic utility of the nuclear gene arginine decarboxylase: an example
A:Reference number: Z16357; MUID:99003705
A:Accession: T08104
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-408 <GAL>
A:Cross-references: EMBL:AF045667; NID:g3421137; PIDN:AAC68512.1; PID:g3421138
C:Genetics:
A:Note: spe2
C:Superfamily: arginine decarboxylase
C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 48.1%; Score 38; DB 2; Length 408;
Best Local Similarity 46.7%; Pred. No. 61;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RTRLRRNHSHKANT 15
Db 107 RAKLRTKHSFGST 121
|:|:|:|:|

Search completed: November 21, 2001, 16:21:54
Job time: 308 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:21:26 ; Search time 17.74 Seconds
(without alignments)
28.965 Million cell updates/sec

Title: US-09-443-986A-1
Perfect score: 79
Sequence: 1 RTRLRNHHSHKANT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	40.5	51.3	576	1 HDAL_STRPU	P56518 strongyloce
2	40	50.6	1089	1 Y553_HUMAN	Q9ukj3 homo sapien
3	39	49.4	926	1 N107_RAT	P52590 rattus norv
4	38	48.1	220	1 RS8_LEINA	P25204 leishmania
5	38	48.1	261	1 CXB6_MOUSE	P70689 mus musculu
6	38	48.1	404	1 YGJK_CAEEL	Q9xuc4 caenorhabdi
7	38	48.1	406	1 SPE2_THECC	O81160 theobroma c
8	38	48.1	407	1 MYC_ASTVU	Q17103 asterias vu
9	38	48.1	502	1 SPE1_LYCES	P49726 lycopersico
10	38	48.1	605	1 RTK2_GEOCY	P42159 geodia cydo
11	38	48.1	702	1 SPE1_ARATH	Q9si64 arabidopsis
12	38	48.1	702	1 SPE1_BRAJU	O82475 brassica ju
13	38	48.1	711	1 SPE2_ARATH	Q33141 arabidopsis
14	38	48.1	725	1 SPE1_DIAAC	Q96412 dianthus ca
15	38	48.1	887	1 SYA_RHIME	P27866 rhizobium m
16	38	48.1	1043	1 TCF8_MESAU	Q60542 mesocricetu
17	38	48.1	1109	1 TCF8_RAT	Q62947 rattus norv
18	38	48.1	1117	1 TCF8_MOUSE	Q64318 mus musculu
19	38	48.1	1124	1 TCF8_HUMAN	P37275 homo sapien
20	38	48.1	1213	1 SMCX_MOUSE	P41230 mus musculu
21	38	48.1	1560	1 SMCX_HUMAN	P41229 homo sapien
22	37	46.8	275	1 LAHL_YEAST	P33399 saccharomyc
23	37	46.8	444	1 YAS2_SCHPO	Q10138 schizosacch
24	37	46.8	572	1 HS60_YEAST	P19882 saccharomyc
25	37	46.8	1018	1 FNBA_STAUV	P14738 staphylococ
26	36	45.6	130	1 CYC3_DESVM	P00132 desulfovibr
27	36	45.6	134	1 RK16_TOBAC	P06384 nicotiana t
28	36	45.6	202	1 VIRE_AGRTU	P15597 agrobacteri
29	36	45.6	256	1 TRYA_DROER	P54624 drosophila
30	36	45.6	308	1 GDF8_HUMAN	Q99988 homo sapien
31	36	45.6	424	1 IRK8_RAT	Q63664 rattus norv
32	36	45.6	446	1 AP50_SCHPO	Q09718 schizosacch
33	36	45.6	476	1 LIPB_PSEFL	P41773 pseudomonas

34	36	45.6	741	1 POL_SOCMV	P15629 soybean chl
35	36	45.6	892	1 HUL4_YEAST	P40985 saccharomyc
36	36	45.6	924	1 CC13_YEAST	P32797 saccharomyc
37	36	45.6	1114	1 TCF8_CHICK	P36197 gallus gall
38	36	45.6	1168	1 LMB3_MOUSE	Q61087 mus musculu
39	36	45.6	1258	1 GLI2_HUMAN	P10070 homo sapien
40	35	44.3	180	1 CYSH_RHITR	O33579 rhizobium t
41	35	44.3	194	1 YF6A_SCHPO	O14256 schizosacch
42	35	44.3	250	1 YK21_YEAST	P36134 saccharomyc
43	35	44.3	256	1 VBR1_BGMV	P06000 bean golden
44	35	44.3	361	1 OTC_ASPTT	Q00291 aspergillus
45	35	44.3	562	1 DHE3_DROME	P54385 drosophila

ALIGNMENTS

```
RESULT 1
HDAL_STRPU STANDARD; PRT; 576 AA.
AC P56518;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HISTONE DEACETYLASE 1 (HD1).
GN HDAC1
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RA Nemer M.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC FAMILY. HD SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF032919; AAB87685.1; -.
CC InterPro; IPR000286; -.
CC InterPro; IPR003084; -.
CC Pfam; PF00850; Hist.deacetyl; 1.
CC PRINTS; PR01270; HDASUPER.
CC PRINTS; PR01271; HSDACETLASE.
CC Hydrolase; Nuclear protein.
CC SEQUENCE 576 AA; 64078 MW; B3D11A844A2088E9 CRC64;
```

Query Match 51.3%; Score 40.5; DB 1; Length 576;
Best Local Similarity 58.8%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Oy 2 TRL-----RRNHSSHA 13

Db 429 TRLPGEGGRDRHRSHKA 445

RESULT 2
Y553_HUMAN

```

ID Y553_HUMAN STANDARD; PRT; 1089 AA.
AC Q9UKJ3; O60300;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN KIA0553.
GN KIA0553.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RS STRAIN=Buffalo; Tissue=Liver;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:31-39(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99408744; PubMed=10477733;
RA Thornton M.A., Poncz M., Korostishevsky M., Yakobson E., Usher S.,
RA Seligsohn U., Peretz H.;
RT "The human platelet alphaIIb gene is not closely linked to its
RT integrin partner beta3";
RL Blood 94:2039-2047(1999).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB011125; BAA25479.1; ALT_INIT.
DR EMBL; AF160252; AAF03681.1; -.
KW Hypothetical protein.
FT DOMAIN 263 266 POLY-LYS.
FT DOMAIN 332 336 POLY-ARG.
FT DOMAIN 359 362 POLY-GLY.
FT DOMAIN 425 430 POLY-GLU.
FT DOMAIN 456 461 POLY-SER.
FT DOMAIN 971 977 POLY-ALA.
FT DOMAIN 1059 1065 POLY-ALA.
FT CONFLICT 207 207 K -> E (IN REF. 1).
FT CONFLICT 267 267 E -> K (IN REF. 1).
FT CONFLICT 278 278 P -> H (IN REF. 1).
FT CONFLICT 846 846 G -> S (IN REF. 1).
SQ SEQUENCE 1089 AA; 117999 MW; 4EF687F9D81A16A5 CRC64;

Query Match 50.6%; Score 40; DB 1; Length 1089;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 RTLRNRHSSHKAN 14
DB 440 RSRGRRHSSHRSS 453

RESULT 3
N107_RAT
ID N107_RAT STANDARD; PRT; 926 AA.
AC P52590;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NUCLEAR PORE COMPLEX PROTEIN NUP107 (NUCLEOPORIN NUP107) (107 KDA
DE NUCLEOPORIN) (P105).

Query Match 50.6%; Score 40; DB 1; Length 1089;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 RTLRNRHSSHKAN 14
DB 440 RSRGRRHSSHRSS 453

RESULT 4
RS8_LEIMA
ID RS8_LEIMA STANDARD; PRT; 220 AA.
AC P25204;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 40S RIBOSOMAL PROTEIN S8.
RN RPS8.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RS STRAIN=LT252;
RX MEDLINE=92158630; PubMed=1741265;
RA Bard E., Janeczko R.;
RT "Sequence and genomic locus of a Leishmania protein homologous to rat
RT ribosomal protein S8.";
RL Nucleic Acids Res. 20:369-369(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X62942; CAA44715.1; -.
DR EMBL; X62942; CAA44714.1; -.
PIR; S20063; S20063.
```

```

GN NUP107.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RS STRAIN=BUFFALO; Tissue=Liver;
RX MEDLINE=94292521; PubMed=8021268;
RA Radu A., Blobel G., Wozniak R.W.;
RT "Nup107 is a novel nuclear pore complex protein that contains a
RT leucine zipper.";
RL J. Biol. Chem. 269:17600-17605(1994).
CC -----
CC -!- FUNCTION: ESSENTIAL COMPONENT OF NUCLEAR PORE COMPLEX.
CC -!- SUBCELLULAR LOCATION: NUCLEAR PORE COMPLEX.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- SIMILARITY: TO YEAST NUCLEOPORIN NUP84.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L31840; AAA74476.1; -.
KW Nuclear protein; Transport.
SQ SEQUENCE 926 AA; 107208 MW; FB80A42C60F4F3B6 CRC64;

Query Match 49.4%; Score 39; DB 1; Length 926;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 TRLRRNHSSHK 12
DB 20 TRTARKHSAHK 30

RESULT 4
RS8_LEIMA
ID RS8_LEIMA STANDARD; PRT; 220 AA.
AC P25204;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 40S RIBOSOMAL PROTEIN S8.
RN RPS8.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RS STRAIN=LT252;
RX MEDLINE=92158630; PubMed=1741265;
RA Bard E., Janeczko R.;
RT "Sequence and genomic locus of a Leishmania protein homologous to rat
RT ribosomal protein S8.";
RL Nucleic Acids Res. 20:369-369(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X62942; CAA44715.1; -.
DR EMBL; X62942; CAA44714.1; -.
PIR; S20063; S20063.
```

```

DR  PIR; S20064; S20064.
DR  InterPro: IPR001047; -.
DR  Pfam: PF01201; Ribosomal_s8e; 1.
DR  PROSITE; PS01193; RIBOSOMAL_S8E; 1.
KW  Ribosomal protein.
SQ  SEQUENCE 220 AA; 24922 MW; 5DD21FE14849445E CRC64;

Query Match 48.1%; Score 38; DB 1; Length 220;
Best Local Similarity 66.7%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TRLRNRHSSHKA 13
   || |||| ||
DB 163 TRRRNRHVEKA 174

RESULT 5
CXB6_MOUSE STANDARD; PRT; 261 AA.
ID CXB6_MOUSE AC F70689;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GAP JUNCTION BETA-6 PROTEIN (CONNEXIN 30) (CX30).
GS GJB6 OR CXN-30.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP RC
STRAIN=C57BL/6;
RX MEDLINE=96279331; PubMed=8663509;
RA Dahl E., Manthey D., Chen Y., Schwarz H.-J., Chang Y.S., Lalley P.A.,
RA Nicholson B.J., Willecke K.;
RT "Molecular cloning and functional expression of mouse connexin-30,a
RT gap junction gene highly expressed in adult brain and skin.";
RL J. Biol. Chem. 271:17903-17910(1996).
CC -!- PAIR OF TRANSMEMBRANE CHANNELS, CONSISTS OF A CLUSTER OF CLOSELY PACKED
CC MATERIALS OF LOW MW DIFFUSE FROM ONE CELL TO A NEIGHBORING CELL.
CC -!- SUBUNIT: A CONNEXON IS COMPOSED OF AN HEXAMER OF CONNEXINS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ADULT BRAIN AND SKIN. LESS
CC EXPRESSION BEFORE BIRTH.
CC -!- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. BETA-TYPE (GROUP I)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z70023; CAA93845.1; -.
DR MGD; MGI:107588; Gjb6.
DR InterPro: IPR000500; -.
DR Pfam; PF00029; connexin; 1.
DR PRINTS; PR00206; CONNEXIN.
DR PROSITE; PS00407; CONNEXINS_1; 1.
DR PROSITE; PS00408; CONNEXINS_2; 1.
KW Gap junction; Transmembrane.
FT DOMAIN 1 19 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 20 40 POTENTIAL.
FT DOMAIN 41 75 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 76 96 POTENTIAL.
FT DOMAIN 97 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 160 POTENTIAL.
FT DOMAIN 161 189 EXTRACELLULAR (POTENTIAL).

```

```
SPE2_THECC
ID SPE2_THECC STANDARD; PRT; 406 AA.
AC O81160;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ARGININE DECARBOXYLASE (EC 4.1.1.19) (ARGDC) (ADC) (FRAGMENT).
GN SPE2.
OS Theobroma cacao (Cacao) (Cocoa).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Malvales; Malvaceae; Theobroma.
OX NCBI_TaxID=3641;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=ADC;
RA Galloway G.L., Malmberg R.L., Price R.A.;
RT "Phylogenetic utility of the nuclear gene arginine decarboxylase: an
RT example from Brassicaceae.";
RL Mol. Biol. Evol. 15:0-0(1998).
CC -!- CATALYTIC ACTIVITY: L-ARGININE = AGMATINE + CO(2).
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE AND MAGNESIUM.
CC -!- PATHWAY: FIRST STEP IN ONE OF THE TWO PUTRESCINE SYNTHESIS
CC PATHWAYS IN PLANTS.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGinine
CC DECARBOXYLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF045666; AAC68511.1; -.
CC Mendel; 31808; Thecc; 1013; 31808.
CC InterPro; IPR000183; -.
CC InterPro; IPR002985; -.
CC Pfam; PF00278; Orn_DAP_Arg_dec; 1.
CC PRINTS; PR01179; ODACRXLASE.
CC PRINTS; PR01180; ARGDCRXLASE.
CC PROSITE; PS00878; ORN_DC_2_1; 1.
CC Putrescine biosynthesis; Spermidine biosynthesis; Lyase; Magnesium;
KW Pyridoxal phosphate; Decarboxylase.
FT NON_TER 1 202 SUBSTRATE-BINDING (BY SIMILARITY).
FT DOMAIN 192 406
FT NON_TER 406
SQ SEQUENCE 406 AA; 44604 MW; A736338EE8ADD4D6 CRC64;

Query Match 48.1%; Score 38; DB 1; Length 406;
Best Local Similarity 46.7%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 RTLRNRHSSHKANT 15
Db 107 RAKLRKHSHGSGT 121
| : : : : : |
| : : : : : |

RESULT 8
MYC_ASTVU
ID MYC_ASTVU STANDARD; PRT; 407 AA.
AC Q17103;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MYC PROTEIN (C-MYC) (FRAGMENT).
GN MYC.
OS Asterias vulgaris (Northern sea star).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC Asteroidea; Forcipulatacea; Forcipulatida; Asteriida; Asterias.
OX NCBI_TaxID=7605;

Query Match 48.1%; Score 38; DB 1; Length 406;
Best Local Similarity 46.7%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 RTLRNRHSSHKANT 15
Db 107 RAKLRKHSHGSGT 121
| : : : : : |
| : : : : : |

RESULT 9
SPE1_LYCES
ID SPE1_LYCES STANDARD; PRT; 502 AA.
AC P49726;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ARGinine DECARBOXYLASE (EC 4.1.1.19) (ARGDC) (ADC).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Pericarp;
RX MEDLINE=94294562; PubMed=8022938;
RA Rastogi R., Dulsion J., Rothstein S.J.;
RT "Cloning of tomato (Lycopersicon esculentum Mill.) arginine
RT decarboxylase gene and its expression during fruit ripening.";
RL Plant Physiol. 103:829-834(1993).
CC -!- CATALYTIC ACTIVITY: L-ARGININE = AGMATINE + CO(2).
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE AND MAGNESIUM.
CC -!- PATHWAY: FIRST STEP IN ONE OF THE TWO PUTRESCINE SYNTHESIS
```

```
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=93026376; PubMed=1408141;
RA Walker C.W., Boom J.D., Marsh A.G.;
RT "First non-vertebrate member of the myc gene family is seasonally
RT expressed in an invertebrate testis.";
RL Oncogene 7:2007-2012(1992).
CC -!- FUNCTION: PARTICIPATES IN THE REGULATION OF GENE TRANSCRIPTION.
CC BINDS DNA BOTH IN A NON-SPECIFIC MANNER AND ALSO SPECIFICALLY TO
CC RECOGNIZES THE CORE SEQUENCE CAC(GA)TG. SEEMS TO ACTIVATE THE
CC TRANSCRIPTION OF GROWTH-RELATED GENES (BY SIMILARITY).
CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M80364; AAA27788.1; -.
CC InterPro; IPR001092; -.
CC InterPro; IPR002418; -.
CC InterPro; IPR003015; -.
CC Pfam; PF00010; HLH; 1.
CC Pfam; PF01056; Myc_N_term; 1.
CC PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
KW Nuclear protein; DNA-binding; Transcription regulation; Activator.
FT NON_TER 1 334
FT DNA_BIND 321 334 BASIC DOMAIN.
FT DOMAIN 335 374 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 380 401 LEUCINE-ZIPPER (POTENTIAL).
SQ SEQUENCE 407 AA; 45673 MW; F4E52DD01182113A CRC64;

Query Match 48.1%; Score 38; DB 1; Length 407;
Best Local Similarity 54.5%; Pred. No. 20;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 RNNHSSHKANT 15
Db 210 KKHSHVHKINT 220
| : : : : : |
| : : : : : |

RESULT 9
SPE1_LYCES
ID SPE1_LYCES STANDARD; PRT; 502 AA.
AC P49726;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ARGinine DECARBOXYLASE (EC 4.1.1.19) (ARGDC) (ADC).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Pericarp;
RX MEDLINE=94294562; PubMed=8022938;
RA Rastogi R., Dulsion J., Rothstein S.J.;
RT "Cloning of tomato (Lycopersicon esculentum Mill.) arginine
RT decarboxylase gene and its expression during fruit ripening.";
RL Plant Physiol. 103:829-834(1993).
CC -!- CATALYTIC ACTIVITY: L-ARGININE = AGMATINE + CO(2).
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE AND MAGNESIUM.
CC -!- PATHWAY: FIRST STEP IN ONE OF THE TWO PUTRESCINE SYNTHESIS
```



```
CC PATHWAYS IN PLANTS.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGinine
CC DECARBOXYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L16582; AAA61347.1; -.
CC InterPro; IPR000183; -.
CC InterPro; IPR002985; -.
CC Pfam; PF00278; Orn_DAP_Arg_dec; 2.
CC PRINTS; PR01179; ODADCRBLXASE.
CC PRINTS; PR01180; ARGDCRBLXASE.
CC PROSITE; PS00878; ODR_DC_2_1; 1.
CC PROSITE; PS00879; ODR_DC_2_2; 1.
CC Pyridoxal biosynthesis; Spermidine biosynthesis; Lyase; Magnesium;
KW Putrescine biosynthesis; Decarboxylase.
FT DOMAIN 226 235 SUBSTRATE-BINDING (BY SIMILARITY).
SQ SEQUENCE 502 AA; 54581 MW; 628D32D4560E9A1F CRC64;

Query Match 48.1%; Score 38; DB 1; Length 502;
Best Local Similarity 46.7%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 RTRLRNHHSHKANT 15
   I : I I I I I I
Db 141 RAKLRTHSGHGSGT 155

RESULT 10
RTK2_GEOCY STANDARD; PRT; 605 AA.
AC P42159;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CLASS II RECEPTOR TYROSINE KINASE (EC 2.7.1.112) (GCTK).
GN TK.
OS Geodia cydonium (Sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC Astrophorida; Geodilidae; Geodia.
OX NCBI_TaxID=6047;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95004895; PubMed=7920862;
RA Schaecke H., Schroeder H.C., Gamulin V., Rinkevich B., Mueller I.M.,
RA Mueller W.E.G.;
RT "Molecular cloning of a tyrosine kinase gene from the marine sponge
RT Geodia cydonium; a new member belonging to the receptor tyrosine
RT kinase class II family."
RL Mol. Membr. Biol. 11:101-107(1994).
RN [2]
RP REVISIONS.
RA Mueller W.E.G.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.
CC -1- PTM: PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X72622; CAA51198.1; -.
CC HSP: P11362; IFCI
CC InterPro; IPR000719; -.
CC InterPro; IPR001245; -.
CC InterPro; IPR002011; -.
CC InterPro; IPR003006; -.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00069; pkinase; 1.
CC PROSITE; PS00107; PROTEIN_KINASE-ATP; FALSE_NEG.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
KW ATP-binding; Phosphorylation.
KW DOMAIN 1 84 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 85 105 POTENTIAL.
FT DOMAIN 106 605 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 346 605 PROTEIN KINASE.
FT NP_BIND 352 360 ATP (BY SIMILARITY).
FT BINDING 393 393 ATP (BY SIMILARITY).
FT ACT_SITE 496 496 BY SIMILARITY.
FT MOD_RES 527 527 PHOSPHORYLATION (AUTO-) (POTENTIAL).
SQ SEQUENCE 605 AA; 67772 MW; E494D0BDBFC9066 CRC64;

Query Match 48.1%; Score 38; DB 1; Length 605;
Best Local Similarity 57.1%; Pred. No. 31;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 TRLRRNHHSHKANT 15
   I : I I I I I I
Db 154 TRLRNRNHIADT 167

RESULT 11
SPEL_ARATH STANDARD; PRT; 702 AA.
AC Q9S164; Q38938;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ARGININE DECARBOXYLASE 1 (EC 4.1.1.19) (ARGDC 1) (ADC-O).
GN SPEL OR AT2G16500 OR FIP15.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; euroids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=96347400; PubMed=8756495;
RA Watson M.B., Malmerberg R.L.;
RT "Regulation of Arabidopsis thaliana (L.) Heynh Arginine decarboxylase
RT by potassium deficiency stress."
RL Plant Physiol. 111:1077-1083(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RL Nature 402:761-768(1999).
```

```
CC -!- CATALYTIC ACTIVITY: L-ARGININE = AGMATINE + CO(2).
CC -!- COPACITOR: PYRIDOXAL PHOSPHATE AND MAGNESIUM.
CC -!- PATHWAY: FIRST STEP IN ONE OF THE TWO PUTRESCINE SYNTHESIS
CC PATHWAYS IN PLANTS.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGinine
CC DECARBOXYLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to licensee@isb-sib.ch).
CC -----
CC EMBL: U52851; AAB09723.1; -
CC EMBL: AC007195; AAD26494.1; -
CC Mende1: 6218; Arath;1013;6218.
CC InterPro: IPR000183; -
CC Pfam: PF00278; Orn.DAP.Arg.dec; 2.
CC PRINTS: PR01179; ODADCRXLASE.
CC PRINTS: PR01180; ARGDCRXLASE.
CC PROSITE: PS00878; ODR_DC_2_1; 1.
CC PROSITE: PS00879; ODR_DC_2_2; 1.
CC Putrescine biosynthesis; Spermidine biosynthesis; Lyase; Magnesium;
CC Pyridoxal phosphate; Decarboxylase.
CC DOMAIN 320 330 SUBSTRATE-BINDING (BY SIMILARITY).
CC CONFLICT 452 452 S -> F (IN REF. 1).
CC SEQUENCE 702 AA; 76175 MW; 8B9D182B2684BCA0 CRC64;

Query Match 48.1%; Score 38; DB 1; Length 702;
Best Local Similarity 46.7%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RTRLRNHHSHKANT 15
Db 235 RAKLRTRKSHGFGST 249
| : || || | : |

RESULT 12
SPEL_BRAJU STANDARD; PRT; 702 AA.
AC O82475;
AT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE ARGinine decarboxylase (EC 4.1.1.19) (ARGDC) (ADC).
GN ADC1.
OS Brassica juncea (Leaf mustard) (Indian mustard).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3707;
RN [1]
RP SEQUENCE FROM N.A.
RA MO H., Pua E.-C.;
RT "Molecular cloning of an arginine decarboxylase cDNA from mustard
RT (Brassica juncea [L.] Czern & Coss).";
RL (In) Plant Gene Register PGR98-160.
CC -!- CATALYTIC ACTIVITY: L-ARGININE = AGMATINE + CO(2).
CC -!- COPACITOR: PYRIDOXAL PHOSPHATE AND MAGNESIUM.
CC -!- PATHWAY: FIRST STEP IN ONE OF THE TWO PUTRESCINE SYNTHESIS
CC PATHWAYS IN PLANTS.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGinine
CC DECARBOXYLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to licensee@isb-sib.ch).
CC -----
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to licensee@isb-sib.ch).
CC -----
CC EMBL: AF077547; AAC62017.1; -
CC Mende1: 33063; Braju;1013;33063.
CC InterPro: IPR002985; -
CC Pfam: PF00278; Orn.DAP.Arg.dec; 2.
CC PRINTS: PR01179; ODADCRXLASE.
CC PRINTS: PR01180; ARGDCRXLASE.
CC PROSITE: PS00878; ODR_DC_2_1; 1.
CC PROSITE: PS00879; ODR_DC_2_2; 1.
CC Putrescine biosynthesis; Spermidine biosynthesis; Lyase; Magnesium;
CC Pyridoxal phosphate; Decarboxylase.
CC DOMAIN 325 335 SUBSTRATE-BINDING (BY SIMILARITY).
CC SEQUENCE 702 AA; 76187 MW; 4FF22C7A8D1B1B92 CRC64;

Query Match 48.1%; Score 38; DB 1; Length 702;
Best Local Similarity 46.7%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RTRLRNHHSHKANT 15
Db 240 RAKLRTRKSHGFGST 254
| : || || | : |

RESULT 13
SPEL_ARATH STANDARD; PRT; 711 AA.
AC O23141;
AT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE ARGinine decarboxylase 2 (EC 4.1.1.19) (ARGDC 2) (ADC 2) (ADC-N).
GN SPE2 OR AT4G34710 OR TAL20.290.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Watson M.B., Yu W., Galloway G., Malmberg R.L.;
RT "Isolation and characterization of a second arginine decarboxylase
RT cDNA from Arabidopsis.";
RL (In) Plant Gene Register PGR97-114.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CV, COLUMBIA;
RA MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Dueterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansoorge W., Brandt P., Grivell L., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohenseil J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Mooljman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Bernesler S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail A., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
RA Pettett A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloembergen M., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
```

RA Neumann S., Argirion A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chefor D., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bieleke C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Groj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Sconeking T., Kallicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Maria M., Martienssen R., McCombie W.R.,
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:769-777(1999).
 CC -i- CATALYTIC ACTIVITY: L-ARGININE = AGMATINE + CO(2).
 CC -i- COFACTOR: PYRIDOXAL PHOSPHATE AND MAGNESIUM.
 CC -i- PATHWAY: FIRST STEP IN ONE OF THE TWO PUTRESCINE SYNTHESIS
 PATHWAYS IN PLANTS.
 CC -i- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGinine
 DECARBOXYLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF009647; AB72179.1; -;
 DR EMBL; AL023094; CAI18850.1; -;
 DR EMBL; AL161586; CAB80188.1; -;
 DR Mendel; 28422; Arath;1013;28422.
 DR InterPro; IPR000183; -;
 DR InterPro; IPR002985; -;
 DR Pfam; PF00278; Orn_DAP_Arg_dec; 2.
 DR PRINTS; PR01179; ODADCRBXLA.
 DR PROSITE; PS00878; ODR_DC_2_1; 1.
 DR PROSITE; PS00879; ODR_DC_2_2; 1.
 KW Putrescine biosynthesis; Spermidine biosynthesis; Lyase; Magnesium;
 KW Pyridoxal phosphate; decarboxylase.
 FT DOMAIN 331 341 SUBSTRATE-BINDING (BY SIMILARITY).
 SQ SEQUENCE 711 AA; 77219 MW; 8E003877173DEE5 CRC64;

Query Match 48.1%; Score 38; DB 1; Length 711;
 Best Local Similarity 46.7%; Pred. No. 37;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 1 RTRLRNHHSHKANT 15
 I : I I I I I I I I
 Db 246 RAKLTKHSGHFGST 260
 RESULT 14
 SPEL_DIACA STANDARD; PRT; 725 AA.
 ID SPEL_DIACA
 AC Q96412;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE ARGinine DECARBOXYLASE (EC 4.1.1.19) (ARGDC) (ADC).
 GN ADC.
 OS Dianthus carophyllus (Carnation) (Clove pink).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;
 OC Caryophyllales; Caryophyllaceae; Dianthus.
 ON NCBI_TaxID=3570;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. WHITE SIM;
 RA Chang K.S., Nam K.H., Lee M.M., Lee S.H., Park K.Y.;
 RT "Nucleotide sequence of cDNA encoding arginine decarboxylase from
 carnation flowers.";
 RL (in) Plant Gene Register PGR96-092.
 CC -i- CATALYTIC ACTIVITY: L-ARGININE = AGMATINE + CO(2).
 CC -i- COFACTOR: PYRIDOXAL PHOSPHATE AND MAGNESIUM.
 CC -i- PATHWAY: FIRST STEP IN ONE OF THE TWO PUTRESCINE SYNTHESIS
 PATHWAYS IN PLANTS.
 CC -i- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGinine
 DECARBOXYLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U63832; AAB67887.1; -;
 DR InterPro; IPR000183; -;
 DR InterPro; IPR002985; -;
 DR Pfam; PF00278; Orn_DAP_Arg_dec; 2.
 DR PRINTS; PR01179; ODADCRBXLA.
 DR PRINTS; PR01180; ARGDCRBXLA.
 DR PROSITE; PS00878; ODR_DC_2_1; 1.
 DR PROSITE; PS00879; ODR_DC_2_2; 1.
 KW Putrescine biosynthesis; Spermidine biosynthesis; Lyase; Magnesium;
 KW Pyridoxal phosphate; decarboxylase.
 FT DOMAIN 340 350 SUBSTRATE-BINDING (BY SIMILARITY).
 SQ SEQUENCE 725 AA; 77676 MW; 6B1C8BA09DE19355 CRC64;
 Query Match 48.1%; Score 38; DB 1; Length 725;
 Best Local Similarity 46.7%; Pred. No. 38;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 1 RTRLRNHHSHKANT 15
 I : I I I I I I I I
 Db 255 RAKLTKHSGHFGST 269
 RESULT 15
 SYA_RHIME
 ID SYA_RHIME STANDARD; PRT; 887 AA.
 AC P27866;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ALANYL-TRNA SYNTHETASE (EC 6.1.1.7) (ALANINE--TRNA LIGASE) (ALARS).
 GN ALAS.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 ON NCBI_TaxID=382;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RA Galibert F., Capela D., Hubler-Barloy F., Gattius M., Batut J.,
 RA Boistard P., Gouzy J., Kahn D., Thebaud P., Goffeau A.,
 RA Purnelle B., Pohl T., Bothe G., Schneider S., Portetelle D.,
 RA Vandenbol M., Puehler A., Becker A., Weidner S.;
 RL Submitted (MAR-2000) to the SWISS-PROT data bank.

```
RN [2]
RP SEQUENCE OF 1-201 FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=91375455; PubMed=1896024;
RA Selbitschka W., Arnold W., Priefer U.B., Rottschliker T.,
RA Schmidt M., Simon R., Puchler A.;
RT "Characterization of recA genes and recA mutants of Rhizobium
RT meliloti and Rhizobium leguminosarum biovar viciae.";
RL Mol. Gen. Genet. 229:86-95(1991).
CC -!- CATALYTIC ACTIVITY: ATP + L-ALANINE + TRNA(ALA) = AMP +
CC PYROPHOSPHATE + L-ALANYL-TRNA(ALA).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X59957; CAA42581.1; -.
DR PIR; S16899; S16899.
DR PIR; S15562; S15562.
DR InterPro; IPR002106; -.
DR InterPro; IPR002318; -.
DR Pfam; PF01411; tRNA-synt_2c; 1.
DR PRINTS; PR00980; TRNASYNTHALA.
DR PROSITE; PS00179; AA_TRNA_LIGASE_II_1; FALSE_NEG.
DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT CONFLICT 135 136 EA -> DT (IN REF. 2).
FT CONFLICT 150 150 R -> H (IN REF. 2).
SQ SEQUENCE 887 AA; 95433 MW; 439810AAFF101A28 CRC64;
```

Query Match 48.1%; Score 38; DB 1; Length 887;
Best Local Similarity 52.9%; Pred. No. 47;
Matches 9; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

Qy 1 RTRLRRNHSS----HKA 13
| ||| ||| : |
Db 557 RARLRANHSATHLLHEA 573

Search completed: November 21, 2001, 16:28:02
Job time: 396 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:20:56 ; Search time 46.09 Seconds
(without alignments)
43.059 Million cell updates/sec

Title: US-09-443-986A-1
Perfect score: 79
Sequence: 1 RTLRNHHSHKANT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_16:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_unclassified:*
 - 13: sp_vertebrate:*
 - 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	43	54.4	517	3	O93870 candida alb
2	43	54.4	1587	10	O9SLQ0 oryza sativ
3	41	51.9	366	10	O9SX36 arabidopsis
4	40	50.6	205	5	O9VLX8 drosophila
5	40	50.6	251	2	O50195 mycobacteri
6	40	50.6	378	5	O16045 drosophila
7	40	50.6	419	3	O9USW3 schizosacch
8	40	50.6	543	5	O9VES0 drosophila
9	40	50.6	1460	4	O75058 homo sapien
10	40	50.6	1486	4	O9UQ08 homo sapien
11	40	50.6	1584	4	O9UQ09 homo sapien
12	39	49.4	144	4	O9NZ91 homo sapien
13	39	49.4	211	4	O9NXL8 homo sapien
14	39	49.4	251	4	O9NZ77 homo sapien
15	39	49.4	381	2	O9ZDS0 rickettsia
16	39	49.4	424	5	O17825 caenorhabdi
17	39	49.4	440	2	O9KYV3 streptomyce
18	39	49.4	551	4	O9H5P2 homo sapien
19	39	49.4	684	2	O51415 pseudomonas

20	39	49.4	757	5	O9V6T1
21	39	49.4	822	5	O61674
22	38	48.1	63	2	O50675
23	38	48.1	117	10	O9XF66
24	38	48.1	175	4	O9NRX2
25	38	48.1	191	2	O9WMP5
26	38	48.1	199	5	O9W2C5
27	38	48.1	258	10	O9LS01
28	38	48.1	261	11	O9RL40
29	38	48.1	309	6	P79352
30	38	48.1	316	6	O9NIY1
31	38	48.1	318	2	O84551
32	38	48.1	401	5	O9U9N9
33	38	48.1	401	10	O81176
34	38	48.1	401	10	O81177
35	38	48.1	403	5	O9XYF8
36	38	48.1	403	10	O81165
37	38	48.1	403	10	O81166
38	38	48.1	403	10	O81180
39	38	48.1	403	10	O81181
40	38	48.1	403	10	O81182
41	38	48.1	403	10	O81183
42	38	48.1	405	10	O81174
43	38	48.1	405	10	O81175
44	38	48.1	405	10	O81178
45	38	48.1	405	10	O81179

ALIGNMENTS

RESULT 1
O93870 PRELIMINARY; PRT; 517 AA.
ID AC O93870;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ZINC CLUSTER TRANSCRIPTION FACTOR FCRLP.
GN FCRL.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B792;
RX MEDLINE=99084962; PubMed=9864335;
RA Talibi D., Raymond M.;
RT "Isolation of a putative Candida albicans transcriptional regulator
involved in pleiotropic drug resistance by functional complementation
of a pdrl pdr3 mutation in Saccharomyces cerevisiae.";
RL J. Bacteriol. 181:231-240(1999).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR CLUSTER
DOMAIN.
CC EMBL; AF057038; AAC98670.1; -.
DR InterPro; IPR001138; -.
DR Pfam; PF00172; Zn_glu; 1.
DR PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR PROSITE; PS00048; ZN2_CY6_FUNGAL_2; 1.
DR SMART; SM00066; GAL4; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Transcription regulation;
KW Zinc.
SQ SEQUENCE 517 AA; 56874 MW; FF71BCE1AED5F21B CRC64;

Query Match 54.4%; Score 43; DB 3; Length 517;
Best Local Similarity 64.3%; Pred. No. 8.3;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 TRLRRNHHSHKANT 15
| | | | | | | | | |

Db 363 TTLRRNSSHSQKT 376

```
RESULT 2
Q3SLQ0
ID Q9SLQ0 PRELIMINARY; PRT: 1587 AA.
AC Q9SLQ0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE GAG-POL POLYPROTEIN (FRAGMENT).
GN GAG-POL.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
OC Oryza.
OX NCBI_TaxID=4530;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=IR36; TRANSPOSON=GYPSY-TYPE RETROTRANSPOSON RIRE7;
RA KumeKawa N., Ohmido N., Fukui K., Ohtsubo E., Ohtsubo H.;
RT "Novel gypsy-type retrotransposon RIRE7: insertion sequence into the
RT tandem repeat sequence trsd which localized in pellicentromeric
RT heterochromatin region in rice.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
CC TRANSCRIPTASE).
CC -!- SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY.
DR EMBL; AB032235; BAA89466.1; -.
DR InterPro; IPR000477; -.
DR InterPro; IPR001584; -.
DR InterPro; IPR001878; -.
DR Pfam; PF00078; rvt; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR Pfam; PF00665; rve; 1.
DR SMART; SM00343; Znf_C2HC; 1.
KW Polypeptide; RNA-directed DNA polymerase; Zinc-finger.
FT NON_TER 1587
SQ SEQUENCE 1587 AA; 180673 MW; 6C2F6BAB2D590969 CRC64;
```

Query Match 54.4%; Score 43; DB 10; Length 1587;
Best Local Similarity 63.6%; Pred. No. 24;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 RNNHSHKANT 15

||||: |||:

Db 92 RNNHSHDANS 102

```
RESULT 3
Q9SX36
ID Q9SX36 PRELIMINARY; PRT: 366 AA.
AC Q9SX36;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE F1413.21 PROTEIN (FRAGMENT).
GN F1413.21.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007980; AAB50048.1; -.
```

FT NON_TER 366 366
SQ SEQUENCE 366 AA; 42039 MW; 81BDB14A36D569B4 CRC64;

Query Match 51.9%; Score 41; DB 10; Length 366;
Best Local Similarity 80.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 LRRNHSHKA 13

||||||| |

Db 47 LRRNHSHGA 56

```
RESULT 4
Q9VLX8
ID Q9VLX8 PRELIMINARY; PRT: 205 AA.
AC Q9VLX8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE CG14538 PROTEIN.
GN CG14538.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokov D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhao Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003618; AAF52551.1; -.
DR Flybase; FBgn0031943; CG14538.
SQ SEQUENCE 205 AA; 22399 MW; ED9DABF040D411AC CRC64;
```

Query Match 50.6%; Score 40; DB 5; Length 205;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRRNHSHKANT 15
 ||| ||| |||
 Db 173 LRRKHSRKRNT 184

RESULT 5
 ID Q50195 PRELIMINARY; PRT; 251 AA.
 AC Q50195;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
 DE L222-ORF6.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97124199; PubMed=8969512;
 RA Fsihi H., De Rossi E., Salazar L., Cantoni R., Labo M., Riccardi G.,
 RA Takiff H.E., Eiglmeyer K., Bergh S., Cole S.T.;
 RT "Gene arrangement and organization in a approximately 76 kb fragment
 encompassing the oric region of the chromosome of Mycobacterium
 leprae";
 RL Microbiology 142:3147-3161(1996).
 DR EMBL: L39923; AAB53126.1; -;
 DR InterPro: IPR000086; -;
 DR Pfam: PF00293; mutf; 1.
 DR PRINTS; PR00502; NUDIXFAMILY.
 DR PROSITE; PS00893; MUTT; 1.
 SQ SEQUENCE 251 AA; 28083 MW; 86F4E04D734219F6 CRC64;

Query Match 50.6%; Score 40; DB 2; Length 251;
 Best Local Similarity 58.3%; Pred. No. 15;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLRRNHSHKA 13
 |::|||::|||
 Db 47 TKPRSHSPRA 58

RESULT 6
 ID O16045 PRELIMINARY; PRT; 378 AA.
 AC O16045;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE C2H2 ZINC FINGER PROTEIN (FRAGMENT).
 GN ANON-EST:FE1E9 OR ANONIE9.
 OS Drosophila yakuba (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97420753; PubMed=92751595;
 RA Schmid K.J., Tautz D.;
 RT "A screen for fast evolving genes from Drosophila.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:9746-9750(1997).
 DR EMBL: AF005848; AAB81481.1; -;
 DR FlyBase; FBgn022416; Dyak\anon-EST:fe1E9.
 DR InterPro; IPR000822; -;
 DR Pfam; PF00096; zf-C2H2; 5.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 3.

DR SMART; SM00355; ZnF_C2H2; 1.
 KW DNA-binding; Metal-binding; Zinc-finger.
 FT NON_TER 1 1
 SQ SEQUENCE 378 AA; 42767 MW; 1E187FEE6D94151E CRC64;

Query Match 50.6%; Score 40; DB 5; Length 378;
 Best Local Similarity 61.5%; Pred. No. 21;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RTRLRRNHSHKA 13
 ||| ||| |||
 Db 269 RTRKRSSSEHS 281

RESULT 7
 ID Q9USW3 PRELIMINARY; PRT; 419 AA.
 AC Q9USW3;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
 DE PUTATIVE GLUCANASE PRECURSOR.
 GN SPAC21B10.07.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972H-;
 RA McDougall R.C., Rajandream M.A., Barrell B.G., Skelton J.,
 RA Churcher C.M.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL121794; CAB57923.1; -;
 SQ SEQUENCE 419 AA; 46508 MW; ABLAD1E14D171FBE CRC64;

Query Match 50.6%; Score 40; DB 3; Length 419;
 Best Local Similarity 53.3%; Pred. No. 24;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 RTRLRRNHSHKANT 15
 ||::||::||::|||
 Db 114 RTAVDRHRPSYKAT 128

RESULT 8
 ID Q9VES0 PRELIMINARY; PRT; 543 AA.
 AC Q9VES0;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE CG10324 PROTEIN.
 GN CG10324.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Bouchon M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cusley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*;"
RL Science 287:2185-2195(2000).
DR EMBL; AE003714; AAF55349.1; -;
DR FlyBase; FBgn0038454; CG10324.
DR InterPro; IPR000467; -;
DR SMART; SM00443; G-patch; 1;
SQ SEQUENCE 543 AA; 61871 MW; 62572907AA6285DA CRC64;

Query Match 50.6%; Score 40; DB 5; Length 543;
Best Local Similarity 63.6%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 5 RRNHSSHKANT 15
|||:|:|
Db 515 RRHSHKHKKT 525

RESULT 9
O75058 PRELIMINARY; PRT; 1460 AA.
AC O75058;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE KIAA0470 PROTEIN.
GN KIAA0470 OR KAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98116662; PubMed=9455484;
RA Seki N., Ohira M., Nagase T., Ishikawa K., Miyajima N., Nakajima D.,
RA Nomura N., Ohara O.;
RT "Characterization of cDNA clones in size-fractionated cDNA libraries
from human brain.";
RL DNA Res. 4:345-349(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Hara Y., Adachi Y.;

RT "Molecular cloning and initial characterization of KAB.;"
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB007939; BAA32315.1; -;
DR EMBL; AB022659; BAA83380.1; -;
DR InterPro; IPR000253; -;
DR Pfam; PF00498; FHA; 1;
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR SMART; SM00240; FHA; 1.
SQ SEQUENCE 1460 AA; 161436 MW; BAD23EBCA19B65F0 CRC64;

Query Match 50.6%; Score 40; DB 4; Length 1460;
Best Local Similarity 77.8%; Pred. No. 75;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 4 LRRNHSSHK 12
|||:|:|
Db 400 LRRHSEHK 408

RESULT 10
O9UQ08 PRELIMINARY; PRT; 1486 AA.
AC O9UQ08;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE KARP-1-BINDING PROTEIN 2 (KAB2).
GN KAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hara Y., Adachi Y.;
RT "Molecular cloning and initial characterization of KAB.;"
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022658; BAA83379.1; -;
DR InterPro; IPR000253; -;
DR Pfam; PF00498; FHA; 1;
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR SMART; SM00240; FHA; 1.
SQ SEQUENCE 1486 AA; 164568 MW; 79312F1C5CB9D04F CRC64;

Query Match 50.6%; Score 40; DB 4; Length 1486;
Best Local Similarity 77.8%; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 4 LRRNHSSHK 12
|||:|:|
Db 400 LRRHSEHK 408

RESULT 11
O9UQ09 PRELIMINARY; PRT; 1584 AA.
AC O9UQ09;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE KARP-1-BINDING PROTEIN 1 (KAB1).
GN KAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hara Y., Adachi Y.;
RT "Molecular cloning and initial characterization of KAB.;"
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB022657; BAA83378.1; -.
DR InterPro; IPR000253; -.
DR Pfam; PF00498; FHA; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR SMART; SM00240; FHA; 1.
SQ SEQUENCE 1584 AA; 175321 MW; A99D76ED374531F0 CRC64;

Query Match 50.6%; Score 40; DB 4; Length 1584;
Best Local Similarity 77.8%; Pred. No. 81;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 LRRNHSHK 12
:|:|:|:|
Db 400 LRRHSEHK 408

RESULT 12
Q9NZ91 ID Q9NZ91 PRELIMINARY; PRT; 144 AA.
AC Q9NZ91;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE UNCHARACTERIZED BONE MARROW PROTEIN BM030.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RA Zhao M., Gu J., Li N., Peng Y., Han Z., Chen Z.;
RT "A novel gene expressed in human bone marrow.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217507; AAF67618.1; -.
DR InterPro; IPR001849; -.
DR Pfam; PF00169; PH; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
SQ SEQUENCE 144 AA; 16521 MW; BFA7E4B2D68D67D9 CRC64;

Query Match 49.4%; Score 39; DB 4; Length 144;
Best Local Similarity 70.0%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLRRNHSH 11
:|:|:|:|
Db 57 SRHRNHSHQ 66

RESULT 13
Q9NRL8 ID Q9NRL8 PRELIMINARY; PRT; 211 AA.
AC Q9NRL8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE BM024.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RA Zhao M., Song H., Li N., Peng Y., Han Z., Chen Z.;
RT "A novel gene expressed in human bone marrow.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF212222; AAF87324.1; -.
DR InterPro; IPR001605; -.
DR InterPro; IPR001849; -.
DR Pfam; PF00169; PH; 1.

DR PRINTS; PR00683; SPECTRINPH.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR SMART; SM00233; PH; 1.
SQ SEQUENCE 211 AA; 23987 MW; 6858B9A4581008ED CRC64;

Query Match 49.4%; Score 39; DB 4; Length 211;
Best Local Similarity 70.0%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLRRNHSH 11
:|:|:|:|
Db 57 SRHRNHSHQ 66

RESULT 14
Q9NZ77 ID Q9NZ77 PRELIMINARY; PRT; 251 AA.
AC Q9NZ77;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE UNCHARACTERIZED BONE MARROW PROTEIN BM046.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RA Zhao M., Gu J., Li N., Peng Y., Han Z., Chen Z.;
RT "A novel gene expressed in human bone marrow.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217522; AAF67633.1; -.
DR InterPro; IPR001605; -.
DR InterPro; IPR001849; -.
DR Pfam; PF00169; PH; 1.
DR PRINTS; PR00683; SPECTRINPH.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR SMART; SM00233; PH; 1.
SQ SEQUENCE 251 AA; 28870 MW; 3C49964F422283C5 CRC64;

Query Match 49.4%; Score 39; DB 4; Length 251;
Best Local Similarity 70.0%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLRRNHSH 11
:|:|:|:|
Db 57 SRHRNHSHQ 66

RESULT 15
Q9ZDS0 ID Q9ZDS0 PRELIMINARY; PRT; 381 AA.
AC Q9ZDS0;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 43.8 KDA PROTEIN.
GN RP255.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; Pubmed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of

RT mitochondria."
RL Nature 396:133-140(1998).
DR EMBL; AJ235271; CAA14717.1; -.
KW Hypothetical protein.
SQ SEQUENCE 381 AA; 43842 MW; E66F1B43FD1D6967 CRC64;

Query Match 49.4%; Score 39; DB 2; Length 381;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 LRRNHSHKAN 14
| | | | |
Db 177 LHYNETSHKLN 187

Search completed: November 21, 2001, 16:27:38
Job time: 402 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:20:49 ; Search time 44.87 Seconds
(without alignments)
21.618 Million cell updates/sec

Title: US-09-443-986A-2
Perfect score: 87
Sequence: 1 GPHRRGRPNRSKRRT 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
A_Geneseq_0601.*			
1:	/SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*		
2:	/SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*		
3:	/SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*		
4:	/SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*		
5:	/SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*		
6:	/SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*		
7:	/SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*		
8:	/SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*		
9:	/SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*		
10:	/SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*		
11:	/SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*		
12:	/SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*		
13:	/SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*		
14:	/SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*		
15:	/SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*		
16:	/SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*		
17:	/SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*		
18:	/SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*		
19:	/SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*		
20:	/SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*		
21:	/SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*		
22:	/SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	87	100.0	16 21 AAB12200
2	79	90.8	17 21 AAB03866
3	53.5	61.5	37 15 AARG0487
4	44	50.6	277 19 AAW81738
5	44	50.6	277 19 AAW64368
6	44	50.6	277 20 AAW39165
7	44	50.6	277 20 AAY39022
8	44	50.6	334 20 AAY36939
9	44	50.6	336 21 AAG46021
10	44	50.6	343 21 AAG46020
11	44	50.6	351 21 AAG46019

12	43	49.4	98	21	AAB58801	Breast and ovarian
13	43	49.4	423	20	AAV03228	Amino acid sequenc
14	43	49.4	423	21	AAB10280	Murine adult thymu
15	42	48.3	104	21	AAB43123	Human ORFX ORF2887
16	42	48.3	107	21	AAG47866	Arabidopsis thalia
17	42	48.3	113	21	AAG47865	Arabidopsis thalia
18	42	48.3	166	21	AAG47864	Arabidopsis thalia
19	42	48.3	232	21	AAG26375	Arabidopsis thalia
20	42	48.3	237	21	AAG26374	Arabidopsis thalia
21	42	48.3	238	21	AAG08621	Arabidopsis thalia
22	42	48.3	238	21	AAG47836	Arabidopsis thalia
23	42	48.3	238	21	AAG47851	Arabidopsis thalia
24	42	48.3	243	21	AAG26373	Arabidopsis thalia
25	42	48.3	244	21	AAG08620	Arabidopsis thalia
26	42	48.3	244	21	AAG47835	Arabidopsis thalia
27	42	48.3	244	21	AAG47850	Arabidopsis thalia
28	42	48.3	248	13	AAR26527	Bovine tp (from cl
29	42	48.3	277	21	AAG47834	Arabidopsis thalia
30	42	48.3	287	21	AAG08619	Arabidopsis thalia
31	42	48.3	288	21	AAG47849	Arabidopsis thalia
32	42	48.3	917	21	AAV53920	A Bcl-2 associated
33	42	48.3	955	21	AAB42231	Human ORFX ORF1995
34	41	47.1	87	20	AAV04818	Mycobacterium spec
35	41	47.1	91	21	AAG34737	Arabidopsis thalia
36	41	47.1	93	21	AAG34735	Arabidopsis thalia
37	41	47.1	102	21	AAB40879	Human ORFX ORF643
38	41	47.1	147	20	AAV04801	Mycobacterium spec
39	41	47.1	184	21	AAB40271	Human ORFX ORF35 p
40	41	47.1	233	20	AAV04815	Mycobacterium spec
41	41	47.1	234	21	AAG27470	Arabidopsis thalia
42	41	47.1	284	21	AAG17964	Arabidopsis thalia
43	41	47.1	285	21	AAG22241	Arabidopsis thalia
44	41	47.1	3066	18	AAW36178	Murine Ataxia-tela
45	40.5	46.6	182	20	AAV29018	T. gondii immunoge

ALIGNMENTS

RESULT 1	
AAB12200	
ID AAB12200 standard; peptide; 16 AA.	
XX	
AC AAB12200;	
XX	
DT 10-NOV-2000 (first entry)	
XX	
DE Gastro-intestinal tract transport receptor targeting agent Zelan145.	
XX	
KW Gastro-intestinal tract; GIT; transport receptor; drug transport;	
KW targeting agent; antibody generation.	
XX	
OS Unidentified.	
XX	
FH Key Location/Qualifiers	
FT Misc-difference 1..16	
FT /note= "D-form residues"	
XX	
PN WO200031546-A1.	
XX	
PD 02-JUN-2000.	
XX	
PF 19-NOV-1999; 99WO-IE00116.	
XX	
PR 19-NOV-1998; 98US-0109036.	
XX	
PA (ELAN-) ELAN CORP PLC.	
XX	
PI O'Mahony DJ, Seveso M;	
XX	
DR WPI; 2000-451689/39.	
XX	
PT Antibody specific to a domain of a gastrointestinal (GIT) targeting	

PT agent such as ZElan033, ZElan088 and ZElan053, useful for detecting,
 XX quantitating and locating GIT targeting agents -
 PS Disclosure; Fig 5; 24pp; English.
 XX
 CC The present sequence is gastro-intestinal tract (GIT) transport receptor
 CC targeting agent, ZElan145. This peptide is capable of facilitating
 CC transport of an active agent e.g. a drug through gastro-intestinal
 CC tissue. The present sequence may be used as an immunogen, for the
 CC generation of antibodies which specifically bind to the present sequence.
 CC The GIT transport receptor targeting agent antibodies are useful for
 CC detecting, quantitating and locating the GIT transport receptor targeting
 CC agents.
 XX
 SQ Sequence 16 AA;

Query Match 100.0%; Score 87; DB 21; Length 16;
 Best Local Similarity 100.0%; Pred. No. 3.8e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPHRRGRPNRSRSKRT 16
 Db 1 gphrrgrpnrsrskrt 16

RESULT 2

AAB03866
 ID AAB03866 standard; peptide; 17 AA.

XX AC AAB03866;
 XX DT 20-OCT-2000 (first entry)

DE GIT receptor targeting peptide ZElan145 (P31 fragment).

XX Retro-inversion peptide; gastrointestinal tract; active agent transport;
 KW GIT; hypertension; diabetes; osteoporosis; haemophilia; anaemia; cancer;
 KW migraine; angina pectoris.

XX Synthetic.

XX Key Location/Qualifiers
 FH Modified-site 1 /note= "Lysine is dansylated"
 FT Misc-difference 2..17 /note= "D-form residue"
 FT

XX WO200031123-A2.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-IE00117.

XX 19-NOV-1998; 98US-0109038.

XX (ELAN-) ELAN CORP PLC.

XX O'Mahony DJ;

XX WPI; 2000-400037/34.

XX Retro-inverted peptide used to deliver active agents across the
 PT gastrointestinal tract to treat hypertension, diabetes, osteoporosis,
 PT haemophilia, anaemia, cancer, migraines and angina pectoris -

XX Disclosure; Page 18; 36pp; English.

XX This invention relates to retro-inverted peptides which specifically bind
 CC to the gastro-intestinal tract receptor hPT1, hPEPT1, D2H or hST. Also
 CC included in the invention are a retro-inverted peptide which enhances the
 CC delivery of an active agent across the gastrointestinal tract (GIT) into
 CC the systemic, portal or hepatic circulation. A composition comprising a

CC retro-inverted peptide bound to a material comprising an active agent
 CC used to treat a mammalian disease or disorder is also disclosed in the
 CC invention. The retro-inversion peptides target gastrointestinal tract
 CC transport receptors to promote in vivo uptake of active agents and/or
 CC enhance active agent delivery across the tract into the systemic
 CC circulation. The gastrointestinal agents (containing retro-inverted
 CC peptides) are used to facilitate the transport of active ingredients
 CC through human or animal gastrointestinal tissue, from the lumen to the
 CC portal, hepatic, or systemic circulation. The compositions containing
 CC these agents can be used to treat or prevent mammalian, especially human,
 CC diseases or disorders, especially hypertension, diabetes, osteoporosis,
 CC haemophilia, anaemia, cancer, migraine, and angina pectoris. The
 CC compositions can be administered in vivo to image selected sites or
 CC tissues, such as the gastrointestinal tract, by using an imaging agent as
 CC the active agent.
 CC The present sequence represents a retro-inversion used in the invention.
 CC The sequence is a P31 16 mer fragment D form retro-inversion peptide.

XX SQ Sequence 17 AA;

Query Match 90.8%; Score 79; DB 21; Length 17;
 Best Local Similarity 93.8%; Pred. No. 6.9e-06;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GPHRRGRPNRSRSKRT 16
 Db 2 gphrrgrpnrsrskrt 17

RESULT 3

AAR60487
 ID AAR60487 standard; peptide; 37 AA.

XX AC AAR60487;

XX DT 10-OCT-1994 (first entry)

DE Peptide analogue of HCV capsid protein used as immunogen.

XX Peptide; retro; invero; retro-inverso; therapy; disease; vaccine;
 KW immunogen; antibody; antibodies; assay; treatment; prophylaxis;
 KW HCV; hepatitis C virus; glycoprotein; epitope; envelope protein;
 KW capsid.

XX Synthetic.

XX Key Location/Qualifiers
 FH Misc-difference 1 /note= "D-form residue"
 FT Misc-difference 2 /note= "D-form residue"
 FT Misc-difference 3 /note= "D-form residue"
 FT Misc-difference 4 /note= "D-form residue"
 FT Misc-difference 5 /note= "D-form residue"
 FT Misc-difference 6 /note= "D-form residue"
 FT Misc-difference 7 /note= "D-form residue"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 9 /note= "D-form residue"
 FT Misc-difference 10 /note= "D-form residue"
 FT Misc-difference 11 /note= "D-form residue"
 FT Misc-difference 12 /note= "D-form residue"
 FT Misc-difference 13 /note= "D-form residue"

FT Misc-difference 14 /note= "D-form residue"
FT Misc-difference 15 /note= "D-form residue"
FT Misc-difference 16 /note= "D-form residue"
FT Misc-difference 17 /note= "D-form residue"
FT Misc-difference 18 /note= "D-form residue"
FT Misc-difference 19 /note= "D-form residue"
FT Misc-difference 20 /note= "D-form residue"
FT Misc-difference 21 /note= "D-form residue"
FT Misc-difference 22 /note= "D-form residue"
FT Misc-difference 23 /note= "D-form residue"
FT Misc-difference 24 /note= "D-form residue"
FT Misc-difference 25 /note= "D-form residue"
FT Misc-difference 26 /note= "D-form residue"
FT Misc-difference 27 /note= "D-form residue"
FT Misc-difference 28 /note= "D-form residue"
FT Misc-difference 29 /note= "D-form residue"
FT Misc-difference 30 /note= "D-form residue"
FT Misc-difference 31 /note= "D-form residue"
FT Misc-difference 32 /note= "D-form residue"
FT Misc-difference 33 /note= "D-form residue"
FT Misc-difference 34 /note= "D-form residue"
FT Misc-difference 35 /note= "D-form residue"
FT Misc-difference 36 /note= "D-form residue"
FT Misc-difference 37 /note= "D-form residue"
FT /note= "D-form residue in amide form"

PN W09405311-A.

XX 17-MAR-1994.

XX 27-AUG-1993; 93WO-AU00441.

XX 27-AUG-1992; 92AU-0004374.

XX (DEAK-) DEAKIN RES LTD.

XX Comis A, Fischer P, Tyler MT;

XX WPI; 1994-100849/12.

XX New retro, inverso and retro-inverso peptide analogues - used as
PT immunogens for the prodn. of antibodies and in diagnosis, therapy
PT and prophylaxis of diseases

XX Example 11; Page 31; 87pp; English.

XX Synthetic peptide antigen analogues of native peptide antigens which
CC are either (1) retro modified, (2) inverso modified or (3) retro-
CC inverso modified with respect to the native antigen can be used to

CC raise antibodies which recognise the native antigen and hence can be
CC used in the treatment and/or prophylaxis of diseases and therapy of
CC disease states. They can be used in the production of vaccines and
CC in assays for antibodies specific for the native antigen. They can
CC have increased stability to degradation after administration and they
CC can be administered orally. This sequence is the retro-inverso
CC version of the sequence described in AAR49986 which is based on the
CC capsid protein of hepatitis C virus, amino acid residues 39-74.

XX Sequence 37 AA;

Query Match 61.5%; Score 53.5; DB 15; Length 37;
Best Local Similarity 75.0%; Pred. No. 0.12;
Matches 12; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 2 PHRRGRPNR-SSKRT 16
| |||| | | |||
Db 12 pgrgrpqrestkrt 27

RESULT 4

AAW81738

ID AAW81738 standard; Protein; 277 AA.

XX AC AAW81738;

XX DT 27-JAN-1999 (first entry)

XX M. tuberculosis immunogenic polypeptide TbH4-Xp1 #2.

XX Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;

XX KW vaccine; pharmaceutical; infection; diagnosis.

XX OS Mycobacterium tuberculosis.

XX PN W09816646-A2.

XX PD 23-APR-1998.

XX PF 07-OCT-1997; 97WO-US18293.

XX PR 13-MAR-1997; 97US-0818112.

XX PR 11-OCT-1996; 96US-0730510.

XX XX (CORI-) CORIXA CORP.

XX PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;

XX PI Reed SG, Skeiky YAW, Twardzik DR, Vedwick TS;

XX XX WPI; 1998-261042/23.

XX DR N-PSDB; AAV64549.

XX Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
XX to develop products for the detection of M. tuberculosis infection
XX and for diagnosis, treatment and prevention of tuberculosis

XX Example 3c; Page 180-181; 230pp; English;

XX This sequence represents an immunogenic portion of a soluble
XX Mycobacterium tuberculosis (MT) antigen which can be used in a method for
XX inducing protective immunity against tuberculosis (TB). This sequence can
XX be formulated into vaccines and/or pharmaceutical compositions for
XX immunising against M. tuberculosis infection or may be used for the
XX diagnosis of tuberculosis.

XX Sequence 277 AA;

Query Match 50.6%; Score 44; DB 19; Length 277;
Best Local Similarity 53.3%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

OY 1 GPHRRGRPNRSRKR 15
   ||||| | : | : |
Db 9 gphrfgapdrgsqrr 23

RESULT 5
AAW64368
ID AAW64368 standard; Protein; 277 AA.
XX AC AAW64368;
XX DT 09-NOV-1998 (first entry)
XX DE Mycobacterium tuberculosis polyprotein.
XX KW Tuberculosis; infection; diagnosis; antigen; TBH-4; XP1.
XX OS Mycobacterium tuberculosis strain H37Rv.
XX OS Mycobacterium tuberculosis strain Erdman.
XX PN W09816645-A2.
XX PD 23-APR-1998.
XX PF 07-OCT-1997; 97WO-US18214.
XX PR 13-MAR-1997; 97US-0818111.
XX PR 11-OCT-1996; 96US-0729622.
XX PA (CORI-) CORIXA CORP.
XX PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;
PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX DR WPI: 1998-251292/22.
XX DR N-PSDB; AAW44440.
XX PT New isolated Mycobacterium tuberculosis polypeptides and DNA - used
PT to develop products for the detection of M. tuberculosis infection
PT and diagnosis of tuberculosis
XX PS Example 3; Page 191-192; 250pp; English.
XX CC This polypeptide is the predicted amino acid sequence deduced from
CC an open reading frame in the reverse complement of Mycobacterium
CC tuberculosis TBH4-XP1 DNA (see AAW44440). The invention relates to
CC methods for diagnosing tuberculosis. It provides polypeptides (see
CC AAW64291-W64379) comprising antigenic or immunogenic portions of M.
CC tuberculosis antigens, as well as DNA sequences encoding such
CC polypeptides, recombinant expression vectors and transformed or
CC transfected host cells. Also claimed are methods and diagnostic
CC kits for detecting M. tuberculosis infection in a patient using
CC these polypeptides, antibodies or oligonucleotide probes and
CC primers.
XX SQ Sequence 277 AA;

Query Match 50.6%; Score 44; DB 19; Length 277;
Best Local Similarity 53.3%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 GPHRRGRPNRSRKR 15
   ||||| | : | : |
Db 9 gphrfgapdrgsqrr 23

RESULT 6
AAAY39165
ID AAY39165 standard; Protein; 277 AA.
XX AC AAY39165;
XX DT 05-NOV-1999 (first entry)
XX DE M. tuberculosis recombinant antigen protein TBH4-XP1(complement).
XX KW Antigen; diagnosis; detection; infection; antibody; immunisation;
XX KW vaccine; immunity.
XX OS Mycobacterium tuberculosis.
XX PN W09942118-A2.
XX PD 26-AUG-1999.

```

```

DT 05-NOV-1999 (first entry)
XX M. tuberculosis antigen TBH4-XP1 reverse complement amino acid sequence.
DE
XX
XX Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW immunotherapy; diagnosis; immunisation; vaccine; infection;
KW immune response; skin test.
XX
OS Mycobacterium tuberculosis.
XX PN W09942076-A2.
XX PD 26-AUG-1999.
XX PF 17-FEB-1999; 99WO-US03268.
XX PR 05-MAY-1998; 98US-0072967.
XX PR 18-FEB-1998; 98US-0025197.
XX PA (CORI-) CORIXA CORP.
XX PI Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX DR WPI: 1999-527409/44.
XX PT New antigens from Mycobacterium tuberculosis useful in diagnostic
XX skin tests and protective or therapeutic vaccines or compositions
XX PS Example 3; Page 175-176; 299pp; English.
XX CC The present invention describes polypeptides comprising an immunogenic
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described
CC are vaccines and fusion protein containing M. tuberculosis Ag's.
CC M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and
CC other polypeptides fragments, can be used in pharmaceutical compositions
CC or vaccines to generate a protective or therapeutic immune response to
CC M. tuberculosis and as reagents in skin tests for diagnosis of
CC tuberculosis. Ag can induce proliferation of, or cytokine secretion
CC by, T, B or natural killer cells and/or macrophages in
CC tuberculosis-immune subjects. AAZ19249 to AAZ19460 and AAY39083 to
CC AAY39225 are used in the exemplification of the present invention.
XX SQ Sequence 277 AA;

Query Match 50.6%; Score 44; DB 20; Length 277;
Best Local Similarity 53.3%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 GPHRRGRPNRSRKR 15
   ||||| | : | : |
Db 9 gphrfgapdrgsqrr 23

RESULT 7
AAAY39022
ID AAY39022 standard; Protein; 277 AA.
XX AC AAY39022;
XX DT 05-NOV-1999 (first entry)
XX DE M. tuberculosis recombinant antigen protein TBH4-XP1(complement).
XX KW Antigen; diagnosis; detection; infection; antibody; immunisation;
XX KW vaccine; immunity.
XX OS Mycobacterium tuberculosis.
XX PN W09942118-A2.
XX PD 26-AUG-1999.

```

```

XX 17-FEB-1999; 99WO-US03265.
XX 05-MAY-1998; 98US-0072596.
XX 18-FEB-1998; 98US-0024753.
XX (CORI-) CORIXA CORP.
XX Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedwick TS;
XX WPI; 1999-527416/44.
XX N-PSDB; AAZ19138.
XX
XX New polypeptide comprising antigenic portions of M. tuberculosis
XX
XX Example 3; Page 221-222; 323pp; English.
XX
XX This invention describes novel recombinant antigens and their encoding
XX nucleic acids derived from Mycobacterium tuberculosis. The novel
XX polypeptides are useful for detecting M. tuberculosis infection in a
XX biological sample by detecting antibodies which bind with the
XX polypeptides, and are useful as vaccines for immunizing against
XX M. tuberculosis infection. The new detection methods are needed as
XX current vaccination strategies do not provide 100% immunity.
XX
XX Sequence 277 AA;
SQ
Query Match 50.6%; Score 44; DB 20; Length 277;
Best Local Similarity 53.3%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 GPHRRGRPNRSRSKR 15
Db ||||| | : | : |
9 gphrfgapdrgsqrr 23
RESULT 8
AAY36939
ID AAY36939 standard; Protein: 334 AA.
XX AC AAY36939;
XX
XX 07-OCT-1999 (first entry)
XX
XX Chlamydia trachomatis surface exposed protein.
XX
XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
XX paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;
XX nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
XX bartholinitis; pneumonia; venereal lymphogranulomatosis.
XX
XX Chlamydia trachomatis.
XX WO9928475-A2.
XX
XX 10-JUN-1999.
XX
XX 27-NOV-1998; 98WO-IB01939.
XX
XX 04-NOV-1998; 98US-0107077.
XX 28-NOV-1997; 97FR-0015041.
XX 17-DEC-1997; 97FR-0016034.
XX
XX (GEST ) GENSET.
XX
XX Griffiths R;
XX
XX WPI; 1999-371125/31.
XX
XX Genome sequence of Chlamydia trachomatis
XX
PS Disclosure; Page 791-792; 1755pp; English.
XX
XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
XX of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
XX vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
XX can also be used to control growth of the microorganism. Chlamydia
XX trachomatis is responsible for a large number of diseases, e.g. eye
XX diseases such as conventional trachoma, nonendemic trachoma,
XX paratrachoma, and inclusion conjunctivitis; genital diseases such as
XX nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
XX perihhepatitis, bartholinitis; pneumonia; venereal lymphogranulomatosis. The polypeptides of the invention
XX may be of use in treating these diseases.
XX
XX Sequence 334 AA;
SQ
Query Match 50.6%; Score 44; DB 20; Length 334;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 2 PHRRGRPNRSRSKR 16
Db ||||| | : | : |
6 phrenpprrmmkrt 20
RESULT 9
AAG46021
ID AAG46021 standard; Protein: 336 AA.
XX AC AAG46021;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 57851.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.

```

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135622.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.8%; Score 44; DB 21; Length 336;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKR 15
| :||||| |||||
Db 121 gekkrgppssskr 135

RESULT 10

AAG46020
ID AAG46020 standard; Protein; 343 AA.

XX AAG46020;

AC AC

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 57850.

XX Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 16-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.

```
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.6%; Score 44; DB 21; Length 343;
Best Local Similarity 60.0%; Pred.No.30;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKR 15
Db 128 gekkrgppgsskr 142
I :||||| |||||
| :||||| |||||

RESULT 11
AAG46019
ID AAG46019 standard; Protein; 351 AA.
XX
AC AAG46019;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 57849.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
```

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.
PR 18-JUN-1999; 99US-0139752.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.6%; Score 44; DB 21; Length 351;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 14
 AAB10280
 ID AAB10280 standard; Protein; 423 AA.
 XX
 AC AAB10280;
 XX
 DT 16-NOV-2000 (first entry)
 XX
 DE Murine adult thymus protein fragment L200_11.
 XX
 KW Secreted protein; cytostatic; immunostimulatory; antimicrobial;
 KW antiviral; immunosuppressive; antiinflammatory; vulnery; cytokine;
 KW cell proliferation; differentiation; regulator; treatment; tumor;
 KW autoimmune disease; inflammatory disorder; wound; microbial infection;
 KW viral disease; graft versus host reaction suppression.
 XX
 OS Mus sp..
 XX
 PN WO200037630-A1.
 XX
 PD 29-JUN-2000.
 XX
 PF 22-DEC-1999; 99WO-US31005.
 XX
 PR 23-DEC-1998; 98US-0220876.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Bowman MR;
 XX
 DR WPI: 2000-442661/38.
 DR N-PSDB; AAA40572.
 XX
 PT Secreted human proteins AS296-11 and AS34-11, useful for treating
 PT tumors, autoimmune diseases, inflammatory disorders, wounds, microbial
 PT infections and viral diseases -
 XX
 PS Disclosure; Page 278-279; 293pp; English.
 XX
 CC This invention describes novel secreted human proteins (I) which have
 CC cytostatic, immunostimulatory, antimicrobial, antiviral,
 CC immunosuppressive, antiinflammatory and vulnery activity and which act
 CC as cytokine, cell proliferation or differentiation regulators. (I)
 CC is useful for treating tumors, autoimmune diseases, inflammatory
 CC disorders, wounds, microbial infections and viral diseases. (I) is also
 CC useful for suppressing graft versus host reaction. AAB10226-B10288
 CC represent the secreted proteins encoded by AAA40490-A40580 which are
 CC described in the method of the invention.
 XX
 SQ Sequence 423 AA;
 Query Match 49.4%; Score 43; DB 21; Length 423;
 Best Local Similarity 66.7%; Pred. No. 52;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 PHRRGRPNRSRS 13
 Db 199 phnsgprstss 210
 || ||| | ||
 RESULT 15
 AAB43123
 ID AAB43123 standard; Protein; 104 AA.
 XX
 AC AAB43123;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF2887 polypeptide sequence SEQ ID NO:5774.
 XX

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 XX thrombosis; contraceptive.
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach M;
 XX
 DR WPI: 2000-602362/57.
 DR N-PSDB; AAC77332.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 4942; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; coagulant; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 104 AA;

Query Match 48.3%; Score 42; DB 21; Length 104;
 Best Local Similarity 60.0%; Pred. No. 20;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 2 PHRRGRPNRSRSKRT 16
 Db 14 pgrnrgsrssrs 28
 | | | | | | | | | |
 | | | | | | | | | |

Search completed: November 21, 2001, 16:20:49
Job time: 299 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:21:20 : Search time 24.83 Seconds
(without alignments)
14.501 Million cell updates/sec

Title: US-09-443-986A-2

Perfect score: 87

Sequence: 1 GPHRRGRPNRSRKR 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	47.1	3066	4	US-08-952-127-12
2	40	46.0	221	2	US-08-943-087-50
3	40	46.0	221	2	US-08-943-087-52
4	40	46.0	221	2	US-08-943-087-54
5	40	46.0	221	2	US-08-943-087-56
6	40	46.0	221	2	US-08-943-087-58
7	40	46.0	221	2	US-08-943-087-60
8	40	46.0	553	2	US-08-943-087-2
9	40	46.0	553	2	US-08-943-087-14
10	40	46.0	553	2	US-08-943-087-16
11	40	46.0	553	2	US-08-943-087-18
12	40	46.0	553	2	US-08-943-087-20
13	40	46.0	553	2	US-08-943-087-22
14	40	46.0	553	2	US-08-943-087-24
15	40	46.0	553	2	US-08-943-087-26
16	40	46.0	553	2	US-08-943-087-28
17	40	46.0	553	2	US-08-943-087-30
18	40	46.0	553	2	US-08-943-087-32
19	40	46.0	553	2	US-08-943-087-34
20	40	46.0	553	2	US-08-943-087-36
21	40	46.0	553	2	US-08-943-087-38
22	40	46.0	553	2	US-08-943-087-40
23	40	46.0	553	2	US-08-943-087-42
24	40	46.0	553	2	US-08-943-087-44
25	40	46.0	553	2	US-08-943-087-46
26	40	46.0	553	2	US-08-943-087-48
27	39	44.8	377	1	US-08-480-882B-6

28	39	44.8	377	1	US-08-480-210-6	Sequence 6, Appli
29	39	44.8	377	1	US-08-220-401-4	Sequence 4, Appli
30	39	44.8	377	2	US-08-437-362-4	Sequence 4, Appli
31	38	43.7	242	2	US-08-845-998-2	Sequence 2, Appli
32	38	43.7	242	4	US-09-206-537-2	Sequence 2, Appli
33	38	43.7	242	4	US-09-430-854-2	Sequence 2, Appli
34	38	43.7	3174	2	US-08-477-451-3	Sequence 3, Appli
35	37.5	43.1	237	1	US-08-910-973-13	Sequence 13, Appli
36	37	42.5	331	4	US-09-025-819-2	Sequence 2, Appli
37	37	42.5	421	2	US-08-484-993B-24	Sequence 24, Appli
38	37	42.5	421	2	US-08-484-158B-24	Sequence 24, Appli
39	37	42.5	421	2	US-08-484-596A-24	Sequence 24, Appli
40	37	42.5	421	2	US-08-480-150A-24	Sequence 24, Appli
41	37	42.5	421	3	US-08-458-731-24	Sequence 24, Appli
42	37	42.5	421	3	US-08-149-223A-24	Sequence 24, Appli
43	37	42.5	614	5	PCT-US95-03236-21	Sequence 21, Appli
44	37	42.5	1708	1	US-08-493-092-2	Sequence 2, Appli
45	37	42.5	1708	1	US-08-508-836A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-952-127-12
: Sequence 12, Application US/08952127
: Patent No. 6211336
: GENERAL INFORMATION:
: APPLICANT: Shiloh, Yosef
: APPLICANT: Tagle, Danilo A.
: APPLICANT: Collins, Francis S.
: TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kohn & Associates
: STREET: 30500 No. 6211336thwestern Hwy., Suite 410
: CITY: Farmington Hills
: STATE: Michigan
: COUNTRY: U.S.
: ZIP: 48334
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/952,127
: FILING DATE:
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Kohn, Kenneth I.
: REGISTRATION NUMBER: 30,995
: REFERENCE/DOCKET NUMBER: 2290.00029
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 810-539-5050
: TELEFAX: 810-539-5055
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3066 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Mus musculus
: US-08-952-127-12

Query Match 47.1%; Score 41; DB 4; Length 3066;
Best Local Similarity 53.3%; Pred. No. 3.4e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKR 15

Db 2796 GAHRRYPNDFSANQ 2810

RESULT 2

US-08-943-087-50
; Sequence 50, Application US/08943087
; Patent No. 5945511
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Kho, Choon J.
; APPLICANT: Jelmeberg, Anna C.
; APPLICANT: Adams, Robyn L.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Farrah, Theresa M.
; TITLE OF INVENTION: CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943.087
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/803.305
; FILING DATE: 20-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: 96-24C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6627
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-943-087-50

Query Match 46.0%; Score 40; DB 2; Length 221;
Best Local Similarity 43.8%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GPHRRGRPNRSRKRT 16
||| : : : ||
Db 195 GPPRAQPSKQART 210

RESULT 3

US-08-943-087-52
; Sequence 52, Application US/08943087
; Patent No. 5945511
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Kho, Choon J.
; APPLICANT: Jelmeberg, Anna C.
; APPLICANT: Adams, Robyn L.

; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Farrah, Theresa M.
; TITLE OF INVENTION: CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943.087
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/803.305
; FILING DATE: 20-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: 96-24C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6627
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-943-087-52

Query Match 46.0%; Score 40; DB 2; Length 221;
Best Local Similarity 43.8%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GPHRRGRPNRSRKRT 16
||| : : : ||
Db 195 GPPRAQPSKQART 210

RESULT 4

US-08-943-087-54
; Sequence 54, Application US/08943087
; Patent No. 5945511
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Kho, Choon J.
; APPLICANT: Jelmeberg, Anna C.
; APPLICANT: Adams, Robyn L.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Farrah, Theresa M.
; TITLE OF INVENTION: CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette


```
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,087
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/803,305
; FILING DATE: 20-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: 96-24C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6627
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-943-087-54

Query Match 46.0%; Score 40; DB 2; Length 221;
Best Local Similarity 43.8%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GPHRRGRNRSRSKRT 16
Db 195 GPPRAQPSKQKQART 210

RESULT 5
US-08-943-087-56
; Sequence 56, Application US/08943087
; Patent No. 5945511
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Kho, Choon J.
; APPLICANT: Jelmlberg, Anna C.
; APPLICANT: Adams, Robyn L.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Fairrah, Theresa M.
; TITLE OF INVENTION: CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,087
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/803,305
; FILING DATE: 20-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: 96-24C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6627
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-943-087-54
```

```
;
; REFERENCE/DOCKET NUMBER: 96-24C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6627
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-943-087-56

Query Match 46.0%; Score 40; DB 2; Length 221;
Best Local Similarity 43.8%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GPHRRGRNRSRSKRT 16
Db 195 GPPRAQPSKQKQART 210

RESULT 6
US-08-943-087-58
; Sequence 58, Application US/08943087
; Patent No. 5945511
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Kho, Choon J.
; APPLICANT: Jelmlberg, Anna C.
; APPLICANT: Adams, Robyn L.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Fairrah, Theresa M.
; TITLE OF INVENTION: CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,087
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/803,305
; FILING DATE: 20-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: 96-24C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6627
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-943-087-58
```

US-08-943-087-58

Query Match 46.0%; Score 40; DB 2; Length 221;
Best Local Similarity 43.8%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GPHRRGRPNRSRSKRT 16
|| || :| : ||
Db 195 GPPRRAPSEKOCART 210

RESULT 7

US-08-943-087-60
; Sequence 60, Application US/08943087
; Patent No. 5945511
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Kho, Choon J.
; APPLICANT: Jelmsberg, Anna C.
; APPLICANT: Adams, Robyn L.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Farrah, Theresa M.
; TITLE OF INVENTION: CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,087
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/803,305
; FILING DATE: 20-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: 96-24C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6627
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-943-087-60

Query Match 46.0%; Score 40; DB 2; Length 221;
Best Local Similarity 43.8%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GPHRRGRPNRSRSKRT 16
|| || :| : ||
Db 195 GPPRRAPSEKOCART 210

RESULT 8

US-08-943-087-2
; Sequence 2, Application US/08943087
; Patent No. 5945511
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Kho, Choon J.
; APPLICANT: Jelmsberg, Anna C.
; APPLICANT: Adams, Robyn L.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Farrah, Theresa M.
; TITLE OF INVENTION: CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,087
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/803,305
; FILING DATE: 20-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: 96-24C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6627
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-943-087-2

Query Match 46.0%; Score 40; DB 2; Length 553;
Best Local Similarity 43.8%; Pred. No. 83;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GPHRRGRPNRSRSKRT 16
|| || :| : ||
Db 224 GPPRRAPSEKOCART 239

RESULT 9

US-08-943-087-14
; Sequence 14, Application US/08943087
; Patent No. 5945511
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Kho, Choon J.
; APPLICANT: Jelmsberg, Anna C.
; APPLICANT: Adams, Robyn L.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Farrah, Theresa M.
; TITLE OF INVENTION: CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:

; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-943-087-18

Query Match 46.0%; Score 40; DB 2; Length 553;
Best Local Similarity 43.8%; Pred. No. 83;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPHRRGRNSRSSKRT 16
|||:|:|
Db 224 GPPRRAQPSKQCART 239

RESULT 12
US-08-943-087-20
; Sequence 20, Application US/08943087
; Patent No. 5945511
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Kho, Choon J.
; APPLICANT: Jelmeberg, Anna C.
; APPLICANT: Adams, Robyn L.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Farrah, Theresa M.
; TITLE OF INVENTION: CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/08/943,087
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/803,305
; FILING DATE: 20-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: 96-24C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6627
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-943-087-20

Query Match 46.0%; Score 40; DB 2; Length 553;
Best Local Similarity 43.8%; Pred. No. 83;

Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 1 GPHRRGRNSRSSKRT 16
|||:|:|
Db 224 GPPRRAQPSKQCART 239

RESULT 13
US-08-943-087-22
; Sequence 22, Application US/08943087
; Patent No. 5945511
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Kho, Choon J.
; APPLICANT: Jelmeberg, Anna C.
; APPLICANT: Adams, Robyn L.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Farrah, Theresa M.
; TITLE OF INVENTION: CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/08/943,087
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/803,305
; FILING DATE: 20-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: 96-24C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6627
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-943-087-22

Query Match 46.0%; Score 40; DB 2; Length 553;
Best Local Similarity 43.8%; Pred. No. 83;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPHRRGRNSRSSKRT 16
|||:|:|
Db 224 GPPRRAQPSKQCART 239

RESULT 14
US-08-943-087-24
; Sequence 24, Application US/08943087
; Patent No. 5945511
; GENERAL INFORMATION:
; APPLICANT: Lok, Si

APPLICANT: Kho, Choon J.
APPLICANT: Jelmsberg, Anna C.
APPLICANT: Adams, Robyn L.
APPLICANT: Whitmore, Theodore E.
APPLICANT: Farrah, Theresa M.
TITLE OF INVENTION: CYTOKINE RECEPTOR
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,087
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/803,305
FILING DATE: 20-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: 96-24C1
TELEPHONE: 206-442-6627
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 553 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-943-087-24

Query Match 46.0%; Score 40; DB 2; Length 553;
Best Local Similarity 43.8%; Pred. No. 83;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKRT 16
|| || |: : ||
Db 224 GPPRAQPSEKQCART 239

RESULT 15
US-08-943-087-26
Sequence 26, Application US/08943087
Patent No. 5945511
GENERAL INFORMATION:
APPLICANT: Lok, Si
APPLICANT: Kho, Choon J.
APPLICANT: Jelmsberg, Anna C.
APPLICANT: Adams, Robyn L.
APPLICANT: Whitmore, Theodore E.
APPLICANT: Farrah, Theresa M.
TITLE OF INVENTION: CYTOKINE RECEPTOR
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA

ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,087
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/803,305
FILING DATE: 20-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: 96-24C1
TELEPHONE: 206-442-6627
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 553 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-943-087-26

Query Match 46.0%; Score 40; DB 2; Length 553;
Best Local Similarity 43.8%; Pred. No. 83;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKRT 16
|| || |: : ||
Db 224 GPPRAQPSEKQCART 239

Search completed: November 21, 2001, 16:21:21
Job time: 310 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:21:54 ; Search time 26.99 Seconds
(without alignments)
45.157 Million cell updates/sec

Title: US-09-443-986A-2

Perfect score: 87
Sequence: 1 GPHRRGRPNRSRKRT 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	52.9	1087	2 T22847	hypothetical prote
2	46	52.9	1260	2 T22845	hypothetical prote
3	44	50.6	156	2 D82799	hypothetical prote
4	44	50.6	229	2 B72700	hypothetical prote
5	44	50.6	296	2 G84747	AT-hook DNA-bindin
6	44	50.6	1122	2 T14180	exit protein - Myc
7	43	49.4	254	1 R3BYS2	ribosomal protein
8	42.5	48.9	1106	2 JQ0405	hypothetical 119.5
9	42.5	48.9	1112	2 S28289	hypothetical prote
10	42	48.3	152	2 F72667	hypothetical prote
11	42	48.3	302	2 A25854	hypothetical prote
12	42	48.3	553	2 T27245	chloramphenicol re
13	42	48.3	659	2 T27246	hypothetical prote
14	41	47.1	282	2 T52466	hypothetical prote
15	41	47.1	284	2 T50673	ribosomal protein
16	41	47.1	448	1 S24756	vicilin-like stora
17	41	47.1	448	1 A60003	nucleocapsid prote
18	41	47.1	590	2 S66956	hypothetical prote
19	41	47.1	1634	2 T26517	hypothetical prote
20	41	47.1	1940	2 F73393	hypothetical prote
21	40	46.0	121	1 R3BS13	ribosomal protein
22	40	46.0	191	2 T28682	hypothetical prote
23	40	46.0	293	2 H70625	hypothetical prote
24	40	46.0	315	2 T41868	hypothetical prote
25	40	46.0	359	2 T14742	hypothetical prote
26	40	46.0	560	2 T23540	hypothetical prote
27	40	46.0	1009	2 T16604	hypothetical prote
28	40	46.0	1638	2 T25352	hypothetical prote
29	39.5	45.4	242	2 F86326	protein F18014.5 [

30 39 44.8 69 2 S39424
31 39 44.8 92 2 S34115
32 39 44.8 101 2 S65494
33 39 44.8 173 2 S62349
34 39 44.8 184 2 F81090
35 39 44.8 341 2 T01538
36 39 44.8 377 1 VHIH79
37 39 44.8 381 1 JQ1725
38 39 44.8 406 2 T36632
39 39 44.8 414 2 T43851
40 39 44.8 468 2 S46791
41 39 44.8 482 2 F71461
42 39 44.8 557 2 B82571
43 39 44.8 716 2 A86181
44 39 44.8 795 2 T34673
45 39 44.8 833 2 A47528

protamine P1 - Aus
sperm-specific pro
sperm-specific pro
l71-3 protein - fr
hypothetical prote
receptor-like prot
nucleocapsid prote
nucleocapsid prote
probable oxidoredu
translation initia
hypothetical prote
hypothetical prote
recombination prot
hypothetical prote
probable SecDF pro
transcription fact

ALIGNMENTS

RESULT 1

T22847

hypothetical protein F57C7.1b - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000

C;Accession: T22847

R;White, S.

submitted to the EMBL Data Library, February 1996

A;Reference number: Z19625

A;Accession: T22847

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1087 <WILL>

A;Cross-references: EMBL:Z69646; PIDN:CAA93475.1; GSPDB:GN000028; CESP:F57C7.1b

A;Experimental source: clone F57C7

C;Genetics:

A;Gene: CESP:F57C7.1b

A;Map position: X

A;Introns: 262/3; 351/3; 391/3; 627/3; 795/2; 957/1

C;Superfamily: bromodomain homology

F;307-364/Domain: bromodomain homology <BROI>

F;579-636/Domain: bromodomain homology <BRO2>

Query Match 52.9%; Score 46; DB 2; Length 1087;

Best Local Similarity 50.0%; Pred. No. 22;

Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKRT 16

DB 417 GPKQGRKSTRGRKKT 432

RESULT 2

T22845

hypothetical protein F57C7.1a - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000

C;Accession: T22845

R;White, S.

submitted to the EMBL Data Library, February 1996

A;Reference number: Z19625

A;Accession: T22845

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1250 <WILL>

A;Cross-references: EMBL:Z69646; PIDN:CAA93473.1; GSPDB:GN000028; CESP:F57C7.1a

A;Experimental source: clone F57C7

C;Genetics:

A;Gene: CESP:F57C7.1a

A;Map position: X

A;Introns: 262/3; 351/3; 391/3; 431/3; 667/3; 835/2; 997/1; 1127/1; 1218/3

C;Superfamily: bromodomain homology
 F;307-364/Domain: bromodomain homology <BRO1>
 F;619-676/Domain: bromodomain homology <BRO2>

Query Match 52.9% Score 46; DB 2; Length 1250;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 8; Conservative 4; Mismatches 0; Gaps 0;

QY 1 GPHRRGRPNRSRKR 16
 || :||| :|| :|| :||
 Db 457 GPKQGRKSTRGRKKT 472

RESULT 3
 D82799
 hypothetical protein XF0483 [imported] - Xylella fastidiosa (strain 9a5c)
 C;Species: Xylella fastidiosa
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C;Accession: D82799
 R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
 Nature 406, 151-157, 2000
 A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A;Reference number: A82515; MUID:20365717
 A;Note: for a complete list of authors see reference number A59328 below
 A;Accession: D82799
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-156 <STM>
 A;Cross-references: GB:AE003898; GB:AE003849; NID:99105329; PIDN:AAF83293.1; GSPDB:GN001
 A;Experimental source: strain 9a5c
 R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000
 A;Authors: Ferreira, V.C.A.; Ferro, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B
 A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A;Reference number: A59328
 A;Contents: annotation
 C;Genetics:
 A;Gene: XF0483

Query Match 50.6% Score 44; DB 2; Length 156;
 Best Local Similarity 53.3%; Pred. No. 8.3;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKR 15
 || :|| :|| :|| :||
 Db 17 GPTKGTGPKCKRTRSKR 31

RESULT 4
 B2700
 hypothetical protein APE1017 - Aeropyrum pernix (strain K1)
 C;Species: Aeropyrum pernix
 C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
 C;Accession: B2700
 R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999
 A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A;Reference number: A72450; MUID:99310339
 A;Accession: B2700
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-229 <RAW>

A;Cross-references: DDBJ:AF000060; NID:95104188; PIDN:BAA80002.1; PID:d1043788; PID:g
 A;Experimental source: strain K1
 C;Genetics:
 A;Gene: APE1017

Query Match 50.6% Score 44; DB 2; Length 229;
 Best Local Similarity 75.0%; Pred. No. 12;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PHRRGRPNRSRS 13
 ||||| |||
 Db 216 PHRRGRPNRSRS 227

RESULT 5
 G84747
 AT-hook DNA-binding protein (AHPL) [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C;Accession: G84747
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487
 A;Accession: G84747
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-296 <STO>
 A;Cross-references: GB:AE002093; NID:g2459442; PIDN:AAB80677.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: At2g33620
 A;Map position: 2

Query Match 50.6% Score 44; DB 2; Length 296;
 Best Local Similarity 60.0%; Pred. No. 14;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKR 15
 | :||| ||||
 Db 81 GEKKRGRPPGSSSKR 95

RESULT 6
 T14180
 exit protein - Mycobacterium smegmatis
 C;Species: Mycobacterium smegmatis
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
 C;Accession: T14180; T14163
 R;Zhu, W.M.; Arceneaux, J.E.L.; Beggs, M.L.; Byers, B.R.; Eisenach, K.D.; Lundrigan,
 Mol. Microbiol. 29, 629-639, 1998
 A;Title: Exochelin genes in Mycobacterium smegmatis: identification of an ABC transpo
 A;Reference number: Z17906; MUID:98385832
 A;Accession: T14180
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1122 <ZHU>
 A;Cross-references: EMBL:AF034152; NID:g3421057; PID:g3421058; PIDN:AAC32046.1
 R;Yu, S.; Fiss, E.; Jacobs Jr., W.R.
 J. Bacteriol. 180, 4676-4685, 1998
 A;Title: Analysis of the exochelin locus in mycobacterium smegmatis: biosynthesis gen
 A;Reference number: Z17898; MUID:98389687
 A;Accession: T14163
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 534-1122 <YU>
 A;Cross-references: EMBL:AF027770; NID:g3560502; PID:g3560505; PIDN:AAC82548.1
 C;Genetics:
 A;Gene: exit
 C;Superfamily: Mycobacterium tuberculosis probable ABC transporter Rv0194; ATP-bindin

Query Match 50.6%; Score 44; DB 2; Length 1122;
 Best Local Similarity 53.3%; Pred. No. 45;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKR 15
 || ||||| : :|
 Db 495 GPARRGRPGAARRR 509

RESULT 7
 R3BYS2
 ribosomal protein S2.e, cytosolic - yeast (Saccharomyces cerevisiae)
 N:Alternate names: omnipotent suppressor protein SUP44; protein YGL123w;
 C:Species: Saccharomyces cerevisiae
 C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Jul-2000
 C:Accession: A36363; A38838; S45501; S64134; S72028; S13309
 R:All-Robyn, J.A.; Brown, N.; Otake, E.; Liebman, S.W.
 Mol. Cell. Biol. 10, 6544-6553, 1990
 A:Title: Sequence and functional similarity between a yeast ribosomal protein and the Es
 A:Reference number: A36363; MUID:91061762
 A:Accession: A36363
 A:Molecule type: DNA
 A:Residues: 1-254 <ALL1>
 A:Cross-references: GB:M59375; EMBL:M38029; NID:g172792; PIDN:AAA63576.1; PID:g172793
 A:Accession: A38838
 A:Molecule type: protein
 A:Residues: 33-55 <ALL2>
 R:Takahara, H.; Tsunasawa, S.; Miyagi, M.; Warner, J.R.
 J. Biol. Chem. 267, 5442-5445, 1992
 A:Title: NH2-terminal acetylation of ribosomal proteins of Saccharomyces cerevisiae.
 A:Reference number: S45501; MUID:92184799
 A:Accession: S45501
 A:Molecule type: protein
 A:Residues: 2-11 <TAK>
 R:Cerdan, E.; Rodriguez-Torres, A.M.; Rodriguez-Belmonte, E.; Tizon, B.; Cadahia, J.L.;
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64134
 A:Accession: S64134
 A:Molecule type: DNA
 A:Residues: 1-254 <GER>
 A:Cross-references: EMBL:Z72645; NID:g1322682; PIDN:CAA96831.1; PID:g1322683; GSPDB:GN00
 A:Experimental source: strain S288C
 R:Lauquin, G.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64122
 A:Accession: S64133
 A:Molecule type: DNA
 A:Residues: 170-254 <LAU>
 A:Cross-references: EMBL:Z72645; GSPDB:GN000007; MIPS:YGL123w
 A:Experimental source: strain S288C
 R:Tizon, B.; Rodriguez-Torres, A.M.; Rodriguez-Belmonte, E.; Cadahia, J.L.; Cerdan, E.
 Yeast 12, 1047-1051, 1996
 A:Title: Identification of a putative methylentetrahydrofolate reductase by sequence an
 A:Reference number: S72026; MUID:97051592
 A:Accession: S72028
 A:Molecule type: DNA
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Residues: 1-254 <TIZ>
 A:Cross-references: EMBL:X94106; NID:g1628448; PIDN:CAA63835.1; PID:g1628451
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
 C:Genetics:
 A:Gene: SGD:SUP44; RP54; MIPS:YGL123w
 A:Cross-references: MIPS:YGL123w; SGD:S0003091
 A:Map position: 7L
 C:Superfamily: Escherichia coli ribosomal protein S5
 C:Keywords: acetylated amino end; blocked amino end; protein biosynthesis; ribosome
 F:2-254/Product: ribosomal protein S2.e #status experimental <MAT>
 F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental

Query Match 49.4%; Score 43; DB 1; Length 254;

Best Local Similarity 50.0%; Pred. No. 18;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKR 16
 | ||||| : :|
 Db 15 GGRNRGRPNRRGPRNT 30

RESULT 8
 JQ0405
 hypothetical 119.5K protein (uvrA region) - Micrococcus luteus
 N:Alternate names: ORF 1 protein
 C:Species: Micrococcus luteus, Micrococcus lysodeikticus
 C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 04-Feb-2000
 C:Accession: JQ0405
 R:Shiota, S.; Nakayama, H.
 Mol. Gen. Genet. 217, 332-340, 1989
 A:Title: Micrococcus luteus homolog of the Escherichia coli uvrA gene: identification
 A:Reference number: S04781; MUID:89364717
 A:Accession: JQ0405
 A:Molecule type: DNA
 A:Residues: 1-1106 <SHI>
 A:Cross-references: EMBL:X15867
 A:Note: this reading frame extends between two stop codons and does not begin with a
 A:Note: the gene encoding this protein overlaps uvrA gene
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo

Query Match 48.9%; Score 42.5; DB 2; Length 1106;
 Best Local Similarity 56.2%; Pred. No. 76;
 Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 GPHRR-GRPNRSRKR 15
 ||||| ||| : :|
 Db 86 GPHRRGRPRAQPEGR 101

RESULT 9
 S28289
 hypothetical protein C38C10.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Jan-2000
 C:Accession: S28289
 R:Thomas, K.
 submitted to the EMBL Data Library, December 1992
 A:Reference number: S28285
 A:Accession: S28289
 A:Molecule type: DNA
 A:Residues: 1-1112 <THO>
 A:Cross-references: EMBL:Z19153
 C:Genetics:
 A:Introns: 412/1; 612/2; 670/3; 676/3; 729/3; 914/2; 984/1
 C:Superfamily: Caenorhabditis elegans hypothetical protein C38C10.5

Query Match 48.9%; Score 42.5; DB 2; Length 1112;
 Best Local Similarity 64.3%; Pred. No. 77;
 Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 2 PHRR-GRPNRSRKR 14
 ||||| : :|
 Db 663 PHRRGEKDRTSK 676

RESULT 10
 F72667
 hypothetical protein APO764 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
 C:Accession: F72667
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
 DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum

A:Reference number: A72450; MUID:99310339

A:Accession: F72667

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-152 <KAT>

A:Cross-references: DBJ:AP000060; NID:g5104188; PID:BAA79742.1; PID:d1043528; PID:g510

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0764

Query Match

Best Local Similarity 48.3%; Score 42; DB 2; Length 152;

Matches 10; Conservative 2; Mismatches 2; Indels 8; Gaps 1;

QY 2 PHRRGRPNSS-----RSSKR 15

Db 70 PHRRGPPSSLPWPPPHRSRPR 91

RESULT 11

A25854

chloramphenicol resistance protein - Escherichia coli plasmid R26

C:Species: Escherichia coli

C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 20-Sep-1999

C:Accession: A25854

R:Dorman, C.J.; Foster, T.J.; Shaw, W.V.

Gene 41, 349-353, 1986

A:Title: Nucleotide sequence of the R26 chloramphenicol resistance determinant and ident

A:Reference number: A25854; MUID:86221720

A:Accession: A25854

A:Molecule type: DNA

A:Residues: 1-302 <DOR>

A:Cross-references: GB:M22614; NID:g151802; PIDN:AAA26079.1; PID:g151803

C:Genetics:

A:Gene: cml

A:Genome: plasmid

C:Superfamily: bicyclomycin resistance protein

C:Keywords: antibiotic resistance

Query Match

Best Local Similarity 48.3%; Score 42; DB 2; Length 302;

Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 PHRRGRPNSSRSKR 15

Db 12 PWSRGPIARSARR 25

RESULT 12

T27245

hypotheical protein Y57G11C.9a - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27245

R:McMurray, A.

submitted to the EMBL Data Library, September 1997

A:Reference number: Z20330

A:Accession: T27245

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-553 <WIL>

A:Cross-references: EMBL:Z99281; PIDN:CAB54457.1; GSPDB:GN00022; CESP:Y57G11C.9a

A:Experimental source: clone Y57G11C

C:Genetics:

A:Gene: CESP:Y57G11C.9a

A:Map position: 4

A:Introns: 67/2; 109/2; 382/1; 418/2; 447/1; 496/3

Query Match

48.3%; Score 42; DB 2; Length 553;

Best Local Similarity 60.0%; Pred. No. 50;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 PHRRGRPNSSRSKRT 16

Db 164 PVRRGRSKRSRSRS 178

RESULT 13

T27246

hypotheical protein Y57G11C.9b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27246

R:McMurray, A.

submitted to the EMBL Data Library, September 1997

A:Reference number: Z20330

A:Accession: T27246

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-659 <WIL>

A:Cross-references: EMBL:Z99281; PIDN:CAB54458.1; GSPDB:GN00022; CESP:Y57G11C.9b

A:Experimental source: clone Y57G11C

C:Genetics:

A:Gene: CESP:Y57G11C.9b

A:Map position: 4

A:Introns: 67/2; 109/2; 382/1; 418/2; 524/2; 553/1; 602/3

Query Match

Best Local Similarity 48.3%; Score 42; DB 2; Length 659;

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 PHRRGRPNSSRSKRT 16

Db 164 PVRRGRSKRSRSRS 178

RESULT 14

T52466

hypotheical protein RF12 [imported] - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 03-Nov-2000

C:Accession: T52466

R:Kato, A.; Suzuki, M.; Kuwahara, A.; Ooe, H.; Higano-Inaba, K.; Kameda, Y.

Gene 239, 309-316, 1999

A:Title: Isolation and analysis of cDNA within a 300 kb Arabidopsis thaliana genomic

A:Reference number: Z25171

A:Accession: T52466

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-282 <KAT>

A:Cross-references: EMBL:AB08017; PIDN:BAA88264.1

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 1

A:Note: RF12

C:Superfamily: Escherichia coli ribosomal protein S5

Query Match

Best Local Similarity 47.1%; Score 41; DB 2; Length 282;

Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPHRRGRPNSSRSKRT 16

Db 34 GPRGRRRGRASEET 49

RESULT 15

T50873

ribosomal protein S2 homolog XM6 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000
C:Accession: T50673
R:Kato, A.; Suzuki, M.; Kuwahara, A.; Ooe, H.; Higano-Inaba, K.; Komeda, Y.
Gene 239, 309-316, 1999
A:Title: Isolation and analysis of cDNA within a 300 kb Arabidopsis thaliana genomic region
A:Reference number: Z25171
A:Accession: T50673
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-284 <KAT>
A:Cross-references: EMBL:AB008016; PIDN:BAA88263.1
A:Experimental source: cultivar Columbia
C:Genetics:
A:Note: XW6
C:Superfamily: Escherichia coli ribosomal protein S5

Query Match 47.1%; Score 41; DB 2; Length 284;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKRT 16
|| ||| |:
Db 36 GPRGRRRGRASEET 51

Search completed: November 21, 2001, 16:21:55
Job time: 309 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:28:02 ; Search time 17.74 Seconds
(without alignments)
30.896 Million cell updates/sec

Title: US-09-443-986A-2

Perfect score: 87

Sequence: 1 GPHRRGRPNRSRSKRT.16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	43	49.4	233	RS2_YEAST	P25443 saccharomyc
2	42.5	48.9	1391	YLD5_CABEL	Q03570 caenorhabdi
3	42	48.3	260	URK1_MOUSE	P52623 mus musculu
4	42	48.3	302	CML_ECOLI	P12056 escherichia
5	41	47.1	223	PRH1_MOUSE	P97458 mus musculu
6	41	47.1	226	PRH1_HUMAN	O75360 homo sapien
7	41	47.1	448	NCAP_CVHOC	P33469 human coron
8	40	46.0	120	RS13_BACSU	P20282 bacillus su
9	40	46.0	392	A2AB_BOVIN	O77700 bos taurus
10	39	44.8	68	HSP1_TACAC	P35311 tachyglossu
11	39	44.8	91	PHI1_MYTED	O04621 mytilus edu
12	39	44.8	145	CRAA_ORYLA	O73919 oryzias lat
13	39	44.8	214	FGF8_CHICK	Q90722 gallus gall
14	39	44.8	377	NCAP_FITPV	P25909 feline infe
15	39	44.8	381	NCAP_CVCAI	P36298 canine ente
16	39	44.8	382	YHH6_YEAST	P32793 saccharomyc
17	39	44.8	391	A2AB_ERIEU	O19012 erinaceus e
18	39	44.8	414	IF2G_HALHA	O93626 halobacteri
19	39	44.8	557	REC_N_XYLFA	O9p602 xylella fas
20	39	44.8	833	MBP1_YEAST	P39678 saccharomyc
21	39	44.8	1972	P531_HUMAN	Q12888 homo sapien
22	38	43.7	79	CMC1_EIMAC	P21959 eimeria ace
23	38	43.7	121	RS13_BACHD	O50632 bacillus ha
24	38	43.7	126	RT12_MARPO	P26871 marchantia
25	38	43.7	184	TRSF_DROSI	Q24669 drosophila
26	38	43.7	193	RL18_TRYBB	P50885 trypanosoma
27	38	43.7	197	TRYP_DROME	P11596 drosophila
28	38	43.7	242	THYL_HUMAN	P20396 homo sapien
29	38	43.7	361	WNT8_BRARE	P51028 brachydanio
30	38	43.7	381	SELP_HUMAN	P49908 homo sapien
31	38	43.7	629	PGMP_BRANA	Q9smm0 brassica na
32	38	43.7	739	DD15_CABEL	Q20875 caenorhabdi
33	38	43.7	832	YFC4_YEAST	P43572 saccharomyc

RESULT 1

ID	RS2_YEAST	STANDARD;	PRT;	253 AA.
AC	P25443;			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	40S RIBOSOMAL PROTEIN S2 (S4) (YS5) (RP12) (OMNIPOTENT SUPPRESSOR			
DE	PROTEIN SUP44).			
GN	RPS2 OR SUP44 OR SUP38 OR YGL123W OR G2893.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91061762; PubMed=2247072;			
RA	All-Robyn J.A., Brown N., Otake E., Liebman S.W.;			
RT	"Sequence and functional similarity between a yeast ribosomal protein			
RT	and the Escherichia coli S5 ram protein.";			
RL	Mol. Cell. Biol. 10:6544-6553(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97051592; PubMed=8896269;			
RA	Tizon B., Rodriguez-Torres A.M., Rodriguez-Belmonte E., Cadahia J.L.,			
RA	Cerdan E.;			
RT	"Identification of a putative methylenetetrahydrofolate reductase by			
RT	sequence analysis of a 6.8 kb DNA fragment of yeast chromosome VII.";			
RL	Yeast 12:1047-1051(1996).			
RN	[3]			
RP	SEQUENCE OF 168-253 FROM N.A.			
RA	Lauquin G.;			
RT	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE OF 1-10.			
RX	MEDLINE=92184799; PubMed=1544921;			
RA	Takakura H., Tsunawasa S., Miyagi M., Warner J.R.;			
RT	"NH2-terminal acetylation of ribosomal proteins of Saccharomyces			
RT	cerevisiae.";			
RL	J. Biol. Chem. 267:5442-5445(1992).			
CC	-/- FUNCTION: IMPORTANT IN THE ASSEMBLY AND FUNCTION OF THE 40S			
CC	RIBOSOMAL SUBUNIT. MUTATIONS IN THIS PROTEIN AFFECTS THE CONTROL			
CC	OF TRANSLATIONAL FIDELITY.			
CC	-/- SIMILARITY: BELONGS TO THE S5P FAMILY OF RIBOSOMAL PROTEINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M59375; AAA63576.1; -			
DR	EMBL; X94106; CAA63835.1; -			
DR	EMBL; Z72645; CAA96831.1; -			

P52958 fusarium so
P15988 gallus gall
P56764 arabidopsis
O14647 homo sapien
Q61496 mus musculu
P02337 clupea pall
P35307 ornithorhyn
P26800 streptomyce
P17527 potato viru
P1757 bacillus st
P30352 gallus gall
Q01130 homo sapien

ALIGNMENTS

DR PIR; A36363; R3BY52.
DR HSSP; P02357; 1PKP.
DR SGD; S0003091; RPS2.
DR InterPro; IPR000851; -.
DR Pfam; PF00333; Ribosomal_S5; 1.
DR PROSITE; PS00585; RIBOSOMAL_S5; 1.
KW Ribosomal protein; Acetylation.
FT INIT_MET 0
FT MOD_RES 1 1 ACETYLATION.
SQ SEQUENCE 253 AA; 27318 MW; 5EC46DC7DE69E6FB CRC64;

Query Match 49.4%; Score 43; DB 1; Length 253;
Best Local Similarity 50.0%; Pred. No. 7;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GPHRGPRNSRSKKT 16
| ||||| | : |
Db 14 GGRNGRPNRRGPRNT 29

RESULT 2
YLD5_CAEEL STANDARD; PRT; 1391 AA.
AC Q03570; P39218; P34581;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 157.0 KDA PROTEIN C38C10.5 IN CHROMOSOME III.
GN C38C10.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]

RP REVISIONS.
RC STRAIN=BRISTOL N2;
RA Jones S.J.M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: C38C10.5A AND C38C10.5B (SHOWN
CC HERE); MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: TO YEAST GR1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; Z29115; CAA82365.1; -.
DR EMBL; Z19153; CAA82365.1; JOINED.
DR EMBL; Z19153; CAA79550.1; -.
DR EMBL; Z29115; CAA79550.1; JOINED.

DR EMBL; Z29115; CAA82366.1; -.
DR EMBL; Z19153; CAA82366.1; JOINED.
DR EMBL; Z19153; CAA79551.1; -.
DR EMBL; Z29115; CAA79551.1; JOINED.
DR PIR; S28289; S28289.
DR WormPep; C38C10.5; CE08648.
KW Hypothetical protein; Alternative splicing.
FT VARSPLIC 671 676 MISSING (IN ISOFORM C38C10.5A).
SQ SEQUENCE 1391 AA; 157165 MW; 8702092D4C55A4BF CRC64;

Query Match 48.9%; Score 42.5; DB 1; Length 1391;
Best Local Similarity 64.3%; Pred. No. 44;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 2 PHRR-GRENSRSK 14
||||| | : ||
Db 663 PHRRGEPKRTSK 676

RESULT 3
URK1_MOUSE STANDARD; PRT; 260 AA.
ID URK1_MOUSE
AC P52623;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE URIDINE KINASE (EC 2.7.1.48) (URIDINE MONOPHOSPHOKINASE) (FRAGMENT).
GN UMPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97108719; PubMed=8951040;
RA Ropp P.A., Traut T.W.;
RT "Cloning and expression of a cDNA encoding uridine kinase from mouse
RT brain.";
RL Arch. Biochem. Biophys. 336:105-112(1996).
CC -!- CATALYTIC ACTIVITY: ATP + URIDINE = ADP + UMP.
CC -!- PATHWAY: PYRIMIDINE SALVAGE PATHWAY.
CC -!- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L31783; AAB50568.1; -.
DR MGD; MGI:98904; Umpk.
DR InterPro; IPR001324; -.
DR Pfam; PF00485; PRK; 1.
KW Transferase; Kinase.
FT NON_TER 1
SQ SEQUENCE 260 AA; 29622 MW; 080F1C82C59C80AC CRC64;

Query Match 48.3%; Score 42; DB 1; Length 260;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 RGRNSRSKKT 16
| | | | : |
Db 223 RGGPNRNRHKT 234

RESULT 4
CML_ECOLI

```

CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U77946; AAB38884.1; -.
CC MGD; MGI:109130; Prop1.
CC InterPro; IPR001356; -.
CC Pfam; PF00046; homeobox; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS50071; HOMEBOX_2; 1.
CC Homeobox; DNA-binding; Nuclear protein.
CC FT DNA_BIND 66 125
CC FT DNAME_BIND 66 125 HOMEBOX.
CC SQ SEQUENCE 223 AA; 25025 MW; CD0AC4DFDC383679 CRC64;

Query Match 47.1%; Score 41; DB 1; Length 223;
Best Local Similarity 61.5%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 RGRPNRSRSKRT 16
Db . :||||:| |
59 QGRPHSRRRRT 71

RESULT 6
PRHL_HUMAN
ID PRHL_HUMAN STANDARD; PRT; 226 AA.
AC Q75360;
DT 15-DEC-1998 (Rel. 37, Created)

```

30-MAY-2000 (REL. 337, Last annotation update)
HOMEOBOX PROTEIN PROPHET OF PIT-1 (PROP-1) (PITUITARY SPECIFIC
HOMEODOMAIN FACTOR).
GN PROP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A., AND VARIANTS CPHD ILE-117 AND CYS-120.
MEDLINE=98122575; PubMed=9462743;
Wu W., Cogan J.D., Pfaeffle R.W., Dasen J.S., Frisch H.,
O'Connell S.M., Flynn S.E., Brown M.R., Mullis P.E., Parks J.S.,
Phillips J.A. III, Rosenfeld M.G.;
"Mutations in PROP1 cause familial combined pituitary hormone
deficiency.";
Nat. Genet. 18:147-149(1998).
[2]
SEQUENCE FROM N.A., AND VARIANT CPHD CYS-73.
MEDLINE=99039899; PubMed=9824293;
Duchesnoy P., Roy A., Dastot F., Ghali I., Teinturier C., Netchine I.,
Cachoux V., Hafez M., Salah N., Chaussain J.L., Goossens M.,
Bouguerès P., Amselem S.;
"Human Prop1: cloning, mapping, genomic structure. Mutations in
familial combined pituitary hormone deficiency.";
FEBS Lett. 437:216-220(1998).
[3]
VARIANT CPHD CYS-120.
MEDLINE=98439627; PubMed=9768691;
Fluck C., Deladoey J., Rutishauser K., Eble A., Marti U., Wu W.,
Mullis P.E.;
"Phenotypic variability in familial combined pituitary hormone
deficiency caused by a PROP1 gene mutation resulting in the
substitution of Arg->Cys at codon 120 (R120C).";
J. Clin. Endocrinol. Metab. 83:3727-3734(1998).
CC -I- FUNCTION: POSSIBLY INVOLVED IN THE ONTOGENESIS OF PITUITARY
CC GONADOTROPHS, AS WELL AS SOMATOTROPHS, LACTOTROPHS AND CAUDOMEDIAL
CC THYROTROPHS.

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN EMBRYONIC PITUITARY.
 CC -!- DISEASE: DEFECTS IN PROP1 ARE A CAUSE OF FAMILIAL COMBINED
 CC PITUITARY HORMONE DEFICIENCY (CPHD) WHICH IS CHARACTERIZED BY
 CC IMPAIRED PRODUCTION OF GROWTH HORMONE (GH) AND ONE OR MORE OF THE
 CC OTHER FIVE ANTERIOR PITUITARY HORMONES.
 CC -!- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF041141; AAC27900.1; -.
 DR EMBL; AF041139; AAC27900.1; JOINED.
 DR EMBL; AF041140; AAC27900.1; JOINED.
 DR EMBL; AF076214; AAC77453.1; -.
 DR EMBL; AF076215; AAC77454.1; -.
 DR MIM; 601538; -.
 DR InterPro; IPR001356; -.
 DR Pfam; PF00046; homeobox; 1.
 DR PROSITE; PS00027; HOMEOBOX_1; 1.
 DR PROSITE; PS00071; HOMEOBOX_2; 1.
 KW Homeobox; DNA-binding; Nuclear protein; Disease mutation.
 FT DNA_BIND 69 128 HOMEOBOX.
 FT VARIANT 73 73 R -> C (IN CPHD).
 FT 117 /FTId=VAR_003768.
 FT 117 F -> I (IN CPHD).
 FT 120 /FTId=VAR_003769.
 FT 120 R -> C (IN CPHD).
 FT 226 /FTId=VAR_003770.
 FT SEQUENCE 226 AA; 24957 MW; D04AYC60F5B07DB4 CRC64;
 SQ

Query Match 47.1%; Score 41; DB 1; Length 226;
 Best Local Similarity 61.5%; Pred. No. 13;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RGRPNRSRSKRT 16
 :|||||
 Db 62 QRGPRHSRRHRT 74

RESULT 7
 NCAP_CVHOC STANDARD; PRT; 448 AA.
 AC P33469;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE NUCLEOCAPSID PROTEIN.
 GN N.
 OS Human coronavirus (strain OC43).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=31631;
 RN [1]
 RP MEDLINE=9243809; PubMed=2541577;
 RX Kanahora T., See L.H., Lai M.M.C.;
 RT "Sequence analysis of nucleocapsid gene and leader RNA of human
 RT coronavirus OC43."
 RL Virus Res. 12:1-9(1989).
 DR PIR; A60003; A60003.
 DR InterPro; IPR01218; -.
 DR Pfam; PF00937; Corona_nucleoca; 1.
 KW Nucleocapsid.
 SQ SEQUENCE 448 AA; 49316 MW; 5193AB1AE0D75626 CRC64;

Query Match 47.1%; Score 41; DB 1; Length 448;
 Best Local Similarity 66.7%; Pred. No. 25;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 RGRPNRSRSKRT 16
 :|||||
 Db 193 RSAPNSRSTSR 204

RESULT 8
 RS13_BACSU STANDARD; PRT; 120 AA.
 AC P20282;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE 30S RIBOSOMAL PROTEIN S13 (BS14).
 GN RPSM.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89213940; PubMed=2496109;
 RA Boylan S.A., Suh J.-W., Thomas S.M., Price C.W.;
 RT "Gene encoding the alpha core subunit of Bacillus subtilis RNA
 RT polymerase is cotranscribed with the genes for initiation factor 1
 RT and ribosomal proteins B, S13, S11, and L17."
 RL J. Bacteriol. 171:2553-2562(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / MAREBURG;
 RX MEDLINE=96186897; PubMed=8635744;
 RA Suh J.W., Boylan S.A., Oh S.H., Price C.W.;
 RT "Genetic and transcriptional organization of the Bacillus subtilis
 RT spc-alpha region."
 RL Gene 169:17-23(1996).
 RN [3]
 RP SEQUENCE OF 85-120 FROM N.A.
 RX MEDLINE=87008431; PubMed=3093467;
 RA Suh J.-W., Boylan S.A., Price C.W.;
 RT "Gene for the alpha subunit of Bacillus subtilis RNA polymerase maps
 RT in the ribosomal protein gene cluster."
 RL J. Bacteriol. 168:65-71(1986).
 CC -!- FUNCTION: INVOLVED IN THE BINDING OF FMET-TRNA AND, HENCE, IN THE
 CC -!- INITIATION OF TRANSLATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE S13P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L47971; AAB06824.1; -.
 DR EMBL; M26414; AAA22215.1; -.
 DR EMBL; M13957; AAA22706.1; -.
 DR EMBL; Z99104; CAB11917.1; -.
 DR PIR; C32307; R3BS13.
 DR Subtilist; BG10730; rpsM.
 DR InterPro; IPR001892; -.
 DR Pfam; PF00416; Ribosomal_S13; 1.
 DR PROSITE; PS00646; RIBOSOMAL_S13; 1.
 KW Ribosomal protein.
 FT INIT_MET 84 84 S -> T (IN REF. 1).
 FT CONFLICT 84 84
 FT SEQUENCE 120 AA; 13656 MW; AD5073E4EA9744EC CRC64;
 SQ

Query Match 46.0%; Score 40; DB 1; Length 120;


```
Best Local Similarity 50.0%; Score 40; DB 1; Length 392;
Matches 9; Conservative 3; Mismatches 2; Indels 4; Gaps 1;

QY 3 HRRGRP---NSRSSKRT 16
    ||||| |||::: ||
Db 90 HRRCLPVRGNSKNNART 107

RESULT 9
A2AB_BOVIN
ID A2AB_BOVIN STANDARD; PRT; 392 AA.
AC O77700;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ALPHA-2B ADRENERGIC RECEPTOR (ALPHA-2B ADRENOCEPTOR) (FRAGMENT).
GN ADRAB2B.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanhope M.J., Madsen O., Waddell V.G., Cleven G.C., de Jong W.W.,
RA Springer M.S.;
RT "Highly congruent molecular support for a diverse superordinal clade
RT of edemic African mammals.;"
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ALPHA-2 ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
CC INDUCED INHIBITION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
CC PROTEINS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y15944; CAA75897.1; -
CC InterPro: IPR000207; -
CC InterPro: IPR000276; -
CC InterPro: IPR002233; -
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODOPSN.
CC PRINTS; PR00559; ADRENRGCA2BR.
CC PRINTS; PR01103; ADRENERGICR.
CC PROSITE; PS00237; G-PROTEIN_RECEP_Fl_1; FALSE_NEG.
CC PROSITE; PSS0262; G-PROTEIN_RECEP_Fl_2; 1.
CC G-protein coupled receptor; Transmembrane; Multigene family.
KW
FT NON_TER 1 1
FT TRANSMEM <1 25 1 (POTENTIAL).
FT DOMAIN 26 36 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 37 62 2 (POTENTIAL).
FT DOMAIN 63 72 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 73 95 3 (POTENTIAL).
FT DOMAIN 96 117 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 118 140 4 (POTENTIAL).
FT DOMAIN 141 156 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 157 180 5 (POTENTIAL).
FT DOMAIN 181 356 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 357 380 6 (POTENTIAL).
FT DOMAIN 381 >392 EXTRACELLULAR (POTENTIAL).
FT DISULFID 72 151 BY SIMILARITY.
FT DOMAIN 283 295 ASP/GLU-RICH (ACIDIC).
FT NON_TER 392 392
SQ SEQUENCE 392 AA; 42723 MW; 6E014510FF47624E CRC64;

Query Match 46.0%; Score 40; DB 1; Length 392;
Best Local Similarity 46.7%; Pred. No. 31;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKR 15
    || : | | | |
Db 193 GPRAKGGGGERSKQ 207

RESULT 10
HSPI_TACAC
ID HSPI_TACAC STANDARD; PRT; 68 AA.
AC P35311;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE SPERM PROTAMINE P1.
GN PRM1 OR PRM-1.
OS Tachyglossus aculeatus aculeatus (Australian echidna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Monotremata; Tachyglossidae; Tachyglossus.
OX NCBI_TaxID=49271;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94094837; PubMed=8269934;
RA Retief J.D., Winkfein R.J., Dixon G.H.;
RT "Evolution of the monotremes. The sequences of the protamine P1 genes
RT of platypus and echidna.;"
RL Eur. J. Biochem. 218:457-461(1993).
CC -1- FUNCTION: PROTAMINES SUBSTITUTES FOR HISTONES IN THE CHROMATIN OF
CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z26848; CAA81444.1; -
CC PIR; S39424; S39424.
CC InterPro: IPR000221; -
CC Pfam; PF00260; protamine_p1; 1.
CC PROSITE; PS00048; PROTAMINE_P1; 1.
CC Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
CC Testis; DNA condensation; Nuclear protein.
FT INIT_MET 0 0 BY SIMILARITY.
SQ SEQUENCE 68 AA; 8673 MW; 41AEF9BB212F7B3 CRC64;

Query Match 44.8%; Score 39; DB 1; Length 68;
Best Local Similarity 66.7%; Pred. No. 8.1;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 RRGPRNRSRKR 15
    |||| : |||: |
Db 50 RRGRRSMRKR 61

RESULT 11
PHIL_MYTED
ID PHIL_MYTED STANDARD; PRT; 91 AA.
AC Q04621;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE SPERM-SPECIFIC PROTEIN PHI-1.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;
```



```
Query Match 44.8%; Score 39; DB 1; Length 214;
Best Local Similarity 52.6%; Pred. No. 25;
Matches 10; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

QY 2 PHRR-----GRPNRSRSK 16
    |||| | : ||||
Db 187 PHRRFEFLNPFNRRSK 205

RESULT 14
NCAP_FIPV
ID NCAP_FIPV STANDARD; PRT; 377 AA.
AC P25909;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE NUCLEOCAPSID PROTEIN (N STRUCTURAL PROTEIN).
GN N.
OS Feline infectious peritonitis virus (strain 79-1146) (FIPV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=33734;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91134997; PubMed=1847259;
RA Vennema H., de Groot R.J., Harbour D.A., Horzinek M.C.,
RA Spaan W.J.M.;
RT "Primary structure of the membrane and nucleocapsid protein genes of
RT feline infectious peritonitis virus and immunogenicity of recombinant
RT vaccinia viruses in kittens.";
RL Virology 181:327-335(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56496; CAA39851.1; -.
DR PIR; B38498; VHIH79.
DR InterPro; IPR001218; -.
DR Pfam; PF00937; Corona_nucleoca; 1.
KW Nucleocapsid.
SQ SEQUENCE 377 AA; 42745 MW; D76382AE5D88D59B CRC64;

Query Match 44.8%; Score 39; DB 1; Length 377;
Best Local Similarity 72.7%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 RRGPRNSRSK 14
    |||| ||| |
Db 17 RRGRSNSGRK 27

RESULT 15
NCAP_CVCAI
ID NCAP_CVCAI STANDARD; PRT; 381 AA.
AC P36298;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE NUCLEOCAPSID PROTEIN.
GN N.
OS Canine enteric coronavirus (strain Insaevc-1) (CCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=36391;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93057357; PubMed=1431811;
```

```
RA Horsburgh B.C., Brierley I., Brown T.D.K.;
RT "Analysis of a 9.6 kb sequence from the 3' end of canine coronavirus
RT genomic RNA.";
RL J. Gen. Virol. 73:2849-2862(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D13096; BAA02414.1; -.
DR PIR; JQ1725; JQ1725.
DR InterPro; IPR001218; -.
DR Pfam; PF00937; Corona_nucleoca; 1.
KW Nucleocapsid.
SQ SEQUENCE 381 AA; 43412 MW; B1E8A0D811D681CB CRC64;

Query Match 44.8%; Score 39; DB 1; Length 381;
Best Local Similarity 72.7%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 RRGPRNSRSK 14
    |||| ||| |
Db 17 RRGRSNSGRK 27

Search completed: November 21, 2001, 16:28:03
Job time: 397 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:27:38 ; Search time 46.09 Seconds
(without alignments)
45.929 Million cell updates/sec

Title: US-09-443-986A-2
Perfect score: 87
Sequence: 1 GPHRRGRPNRSRKRT 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- SPTREMBL_16:*
- 1: sp.archaea:*
 - 2: sp.bacteria:*
 - 3: sp.fungi:*
 - 4: sp.human:*
 - 5: sp.invertebrate:*
 - 6: sp.mammal:*
 - 7: sp.mhc:*
 - 8: sp.organelle:*
 - 9: sp.phage:*
 - 10: sp.plant:*
 - 11: sp.rodent:*
 - 12: sp.unclassified:*
 - 13: sp.vertebrate:*
 - 14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	49	56.3	4498	13 Q93291	Q93291 fuğu rubrip
2	48	55.2	753	10 Q9SDD2	Q9sdd2 oryza sativ
3	46	52.9	1087	5 Q20948	Q20948 caenorhabdi
4	46	52.9	1250	5 Q20947	Q20947 caenorhabdi
5	44	50.6	156	2 Q9PG19	Q9pg19 xylella fas
6	44	50.6	191	5 Q26468	Q26468 schistocerc
7	44	50.6	229	1 Q9VD95	Q9vd95 aeropyrum p
8	44	50.6	296	10 Q22812	Q22812 arabidopsis
9	44	50.6	351	10 Q23142	Q23142 arabidopsis
10	44	50.6	724	14 Q9Q3G5	Q9q3g5 turkey astr
11	44	50.6	1122	2 Q85018	Q85018 mycobacteri
12	43	49.4	169	11 Q9WV90	Q9wv90 marimota mon
13	43	49.4	409	14 Q64961	Q64961 avian infec
14	43	49.4	409	14 Q64931	Q64931 avian infec
15	43	49.4	409	14 Q9J4A3	Q9j4a3 avian infec
16	43	49.4	453	5 Q9V537	Q9v537 drosophila
17	43	49.4	693	5 Q9VNG4	Q9vng4 drosophila
18	42	48.3	152	1 Q9YE05	Q9ye05 aeropyrum p
19	42	48.3	193	2 Q9K526	Q9k526 mycobacteri

ALIGNMENTS

RESULT 1

ID	Q93291	PRELIMINARY;	PRT; 4498 AA.
AC	Q93291;		
DT	01-NOV-1998 (TReMBLrel. 08, Created)		
DT	01-NOV-1998 (TReMBLrel. 08, Last sequence update)		
DT	01-MAR-2001 (TReMBLrel. 16, Last annotation update)		
DE	MLL.		
GN	MLL.		
OS	Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;		
OC	Tetraodontidae; Takifugu.		
OX	NCBI_TaxID=31033;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=98345146; PubMed=9681821;		
RA	Caldas C., Kim M.-H., MacGregor A., Cain D., Aparicio S.,		
RA	Wiedemann L.N.;		
RT	"Isolation and characterization of a pufferfish MLL (mixed lineage		
RT	leukemia)-like gene (fml1) reveals evolutionary conservation in		
RT	vertebrate genes related to Drosophila trithorax.";		
RL	OncoGene 16:3233-3241(1998).		
DR	EMBL: AF036382: AAC41377.1; -.		
DR	InterPro: IPR001005; -.		
DR	InterPro: IPR001214; -.		
DR	InterPro: IPR001965; -.		
DR	InterPro: IPR002857; -.		
DR	InterPro: IPR003616; -.		
DR	Pfam: PF00628; PHD; 3.		
DR	Pfam: PF00856; SET; 1.		
DR	Pfam: PF02008; zf-CXXC; 1.		
DR	PROSITE; PS00037; MYB_1; UNKNOWN_1.		
DR	SMART; SM00508; PostSET; 1.		
SQ	SEQUENCE 4498 AA; 486562 MW; 7B6A7ECFFA5EDFCB CRC64;		

Query Match 56.3%; Score 49; DB 13; Length 4498;
Best Local Similarity 62.5%; Pred. No. 20;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 GPHRRGRPNRSRKRKT 16
  | | | | | | | | | |
Db 168 GEKRRGRPPSSSKKS 183

RESULT 2
Q9SDD2 PRELIMINARY: PRT; 753 AA.
AC Q9SDD2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE SIMILAR TO ARABIDOPSIS THALIANA DNA CHROMOSOME 4.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
OC Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
  clone:PO011D01.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP0000969; BAA88534.1; -.
DR InterPro; IPR002052; -.
DR PROSITE; PS00092; N6.MTASE; UNKNOWN_1.
SQ SEQUENCE 753 AA; 83910 MW; EB1BCDD2DC7E080A CRC64;

Query Match 55.2%; Score 48; DB 10; Length 753;
Best Local Similarity 81.8%; Pred. No. 6.1;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 RGRPNRSKR 15
  | | | | | | | |
Db 286 RGRPNRSQR 296

RESULT 3
Q20948 PRELIMINARY: PRT; 1087 AA.
AC Q20948;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE F57C7.1B PROTEIN.
GN F57C7.1B.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA White S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
  Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
  Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
  Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
  Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
  Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
  Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
  Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
  Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
  Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
  "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
  elegans.";
RL Nature 368:32-38(1994).
DR EMBL; 269646; CAA93473.1; -.
DR InterPro; IPR00130; -.
DR Pfam; PF00439; bromodomain; 3.
DR PRINTS; PR00503; BROMODOMAIN.
DR PROSITE; PS00633; BROMODOMAIN_1; UNKNOWN_1.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SMART; SM00297; BROMO; 1.
SQ SEQUENCE 1250 AA; 138569 MW; FA2EFF772A8E3DE0F CRC64;

Query Match 52.9%; Score 46; DB 5; Length 1250;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRKRKT 16
  | | | | | | | | | |
Db 457 GPKQGRKSTRGRKKT 472
```

```

RESULT 5
Q9PG19 ID Q9PG19 PRELIMINARY; PRT; 156 AA.
AC Q9PG19
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL PROTEIN XF0483.
GN XF0483.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Carraro L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Franco M.C., Frohme M., Furian L.R.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kurame E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.F.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peloto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.A., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 405:151-159(2000)
DR EMBL; AE003898; AAF83293.1; -
KW Hypothetical protein.
SQ SEQUENCE 156 AA; 17837 MW; 119CDC20E476EE7D CRC64;

Query Match 50.68; Score 44; DB 2; Length 156;
Best Local Similarity 53.38; Pred. No. 7.1;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GPHRRGRPNRSRSKR 15
II : I : I : I : I :
Db 17 GPTKTRKPKRTSKR 31

RESULT 6
Q26468 ID Q26468 PRELIMINARY; PRT; 191 AA.
AC Q26468
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE TRANSFORMING GROWTH FACTOR-BETA (FRAGMENT).
OS Schistocerca americana (American grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridoidea; Acridoidea; Acrididae; Schistocerca.
OX NCBI_TaxID=7009;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA80002.1; -
DR InterPro; IPR002220; -
DR ProDom; PD001859; -; 1.

Query Match 50.68; Score 44; DB 5; Length 191;
Best Local Similarity 57.98; Pred. No. 8.4;
Matches 11; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

Qy 1 GPHRR---GRPNRSRSKR 15
II : I : I : I : I :
Db 65 GQRRSRKNGAPPSSRSKR 83

RESULT 7
Q9YD95 ID Q9YD95 PRELIMINARY; PRT; 229 AA.
AC Q9YD95;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 24.1 KDA PROTEIN APE1017.
GN APE1017.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA80002.1; -
DR InterPro; IPR002220; -
DR ProDom; PD001859; -; 1.

```

```
KW Hypothetical protein.
SQ SEQUENCE 229 AA; 24065 MW; 265A6CF47121AF09 CRC64;

Query Match 50.6%; Score 44; DB 1; Length 229;
Best Local Similarity 75.0%; Pred. No. 9.9;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PHRRGRPNRSRS 13
   ||||| |||
DB 216 PHRRGRHPSRHS 227

RESULT 8
O22812 PRELIMINARY; PRT; 296 AA.
AC O22812;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE PUTATIVE DNA-BINDING PROTEIN PD1.
GN F4P9.39.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D., Somerville C.R.,
RA Venter J.C.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC002332; AAB0677.1;...
DR Mendel; 25331; Arath; 2835; 25331.
DR InterPro; IPR000637; -.
DR Pfam; PF02178; AT_hook; 2.
DR SMART; SM00384; AT_hook; 1.
KW DNA-binding.
SQ SEQUENCE 296 AA; 30296 MW; A30EB93F3C408C1E CRC64;

Query Match 50.6%; Score 44; DB 10; Length 296;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRSKR 15
   | :|||| ||||
DB 81 GEKRGRRPPGSSSKR 95

RESULT 9
O23142 PRELIMINARY; PRT; 351 AA.
AC O23142;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE AT-HOOK PROTEIN 1.
GN AHPL.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FLOWER BUDS;
RA Hofmann W.A., Saedler H., Huijser P.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ222585; CAA10857.1; -.
DR Mendel; 25479; Arath; 2835; 25479.

DR InterPro; IPR000637; -.
DR Pfam; PF02178; AT_hook; 2.
DR SMART; SM00384; AT_hook; 1.
SQ SEQUENCE 351 AA; 36317 MW; 013B1B90EEF147A9 CRC64;

Query Match 50.6%; Score 44; DB 10; Length 351;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPHRRGRPNRSRSKR 15
   | :|||| ||||
DB 136 GEKRGRRPPGSSSKR 150

RESULT 10
Q903G5 PRELIMINARY; PRT; 724 AA.
AC Q903G5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CAPSID PROTEIN.
OS turkey astrovirus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Astroviridae;
OC Astrovirus.
OX NCBI_TaxID=70793;
RN [1]
RP SEQUENCE FROM N.A.
RA Kocil M.D., Seal B.S., Schultz-Cherry S.;
RT "Identification and molecular characterization of turkey
RT astrovirus.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF206663; AAF18454.1; -.
SQ SEQUENCE 724 AA; 80265 MW; DB183AC7885EDEF4 CRC64;

Query Match 50.6%; Score 44; DB 14; Length 724;
Best Local Similarity 69.2%; Pred. No. 27;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 RRGPRNSRSKRKT 16
   |||| |||| |
DB 17 RRGPRNSRSRSRS 29

RESULT 11
O85018 PRELIMINARY; PRT; 1122 AA.
AC O85018;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE EXIT.
GN EXIT.
OS Mycobacterium smegmatis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LR222;
RX MEDLINE=98385832; PubMed=9720878;
RA Zhu W.M., Arceneaux J.E.L., Beggs M.L., Byers B.R., Eisenach K.D.,
RA Lundrigan M.D.;
RT "Exochelin genes in Mycobacterium smegmatis: identification of an ABC
RT transporter and two non-ribosomal peptide synthetase genes.";
RL Mol. Microbiol. 29:629-639(1998).
RN [2]
RP SEQUENCE OF 534-1122 FROM N.A.
RC STRAIN=MC2155;
RX MEDLINE=95191405; PubMed=7885234;
RA Fiss E.H., Yu S., Jacobs W.R. Jr.;
```


RT "Identification of genes involved in the sequestration of iron in
 RL mycobacteria: the ferric exochelin biosynthetic and uptake pathways.",
 RN [3]
 RP SEQUENCE OF 534-1122 FROM N.A.

RC STRAIN-MC2155;
 RX MEDLINE=98389687; PubMed=9721311;
 RA Yu S., Fiss E., Jacobs W.R. Jr.;
 RT "Analysis of the exochelin locus in Mycobacterium smegmatis:
 RT biosynthesis genes have homology with genes of the peptide synthetase
 RT family.";

RL J. Bacteriol. 180:4676-4685(1998).

DR EMBL: AF034152; AAC32046.1; -;

DR EMBL: AF027770; AAC82548.1; -;

DR InterPro: IPR001140; -;

DR InterPro: IPR001687; -;

DR InterPro: IPR003439; -;

DR InterPro: IPR003593; -;

DR Pfam: PF00005; ABC_tran; 1.

DR Pfam: PF00664; ABC_membrane; 1.

DR SMART: SM00382; AAA; 1.

KW Hypothetical protein.

SQ SEQUENCE 1122 AA; 120386 MW; 4EC826C67914A57A CRC64;

Query Match 50.6%; Score 44; DB 2; Length 1122;

Best Local Similarity 53.3%; Pred. No. 40;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GPHRRGRPNRSRKR 15

Db 495 GPARRGREGAARHR 509

||| ||||| : ||

RESULT 12

ID Q9WV90 PRELIMINARY; PRT; 169 AA.

AC Q9WV90;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

DE FAS LIGAND (FRAGMENT).

OS Marmota monax (Woodchuck).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;

OC Marmota.

OX NCBI_TaxID=9995;

FN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=HEALTHY LIVER;

RA Hodgson P.D., Grant M.D., Michalak T.I.;

RT "Perforin and Fas/Fas ligand-mediated cytotoxicity in acute and

RL Clin. Exp. Immunol. 0:0-0(1999).

DR EMBL: AF152368; AAD38387.1; -;

DR InterPro: IPR000478; -;

DR Pfam: PF00229; TNF; 1.

DR PROSITE; PS00251; TNF_1; 1.

DR PROSITE; PS00049; TNF_2; 1.

DR SMART; SM00207; TNF; 1.

FT NON_TER 169 169

SQ SEQUENCE 169 AA; 19274 MW; FDE395B014717B6B CRC64;

Query Match 49.4%; Score 43; DB 11; Length 169;

Best Local Similarity 72.7%; Pred. No. 11;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 HRRGRPNRSR 13

Db 45 HLTGKPNRSR 55

|||||

RESULT 13

Q64961

ID Q64961 PRELIMINARY; PRT; 409 AA.

AC Q64961;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE NUCLEOCAPSID PROTEIN.

OS avian infectious bronchitis virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;

OC Coronaviridae; Coronavirus.

OX NCBI_TaxID=11120;

FN [1]

RP SEQUENCE FROM N.A.

RA Williams A.K., Wang L., Sneed L.S., Collisson E.W.;

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: M85246; AAA91857.1; -;

DR InterPro: IPR001218; -;

DR Pfam: PF00937; Corona_nucleoca; 1.

KW Nucleocapsid.

SQ SEQUENCE 409 AA; 45091 MW; FA7DB0F7E406C959 CRC64;

Query Match 49.4%; Score 43; DB 14; Length 409;

Best Local Similarity 53.3%; Pred. No. 24;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 PHRRGRPNRSRKR 16

Db 336 PRPKSRPNRPATRT 350

||||| : ||

RESULT 14

Q64931

ID Q64931 PRELIMINARY; PRT; 409 AA.

AC Q64931; O12850; O12563;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE NUCLEOCAPSID PROTEIN.

GN GENE 6.

OS avian infectious bronchitis virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;

OC Coronaviridae; Coronavirus.

OX NCBI_TaxID=11120;

FN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CU-T2;

RA Jia W., Karaca K., Parrish C.R., Naqi S.A.;

RT "A novel variant of avian infectious bronchitis virus resulting from

RL recombination among three different strains.";

Arch. Virol. 140:259-271(1995).

FN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CU-T2;

RA Jia W., Naqi S.A.;

RT "Sequence analysis of gene 3, gene 4 and gene 5 of avian infectious

RL bronchitis virus strain CU-T2.";

Gene 189:189-193(1997).

DR EMBL: U04805; AAC54068.1; -;

DR EMBL: U49858; AAC57080.1; -;

DR InterPro: IPR001218; -;

DR Pfam: PF00937; Corona_nucleoca; 1.

KW Nucleocapsid.

SQ SEQUENCE 409 AA; 45027 MW; F802296275BFF5D7 CRC64;

Query Match 49.4%; Score 43; DB 14; Length 409;

Best Local Similarity 53.3%; Pred. No. 24;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 PHRRGRPNRSRSKRT 16
 | : |||| : ||
 Db 336 PRPKSRPNRSRPATRT 350

RESULT 15

Q9J4A3
 ID Q9J4A3 PRELIMINARY; PRT; 409 AA.
 AC Q9J4A3;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE NUCLEOCAPSID.
 GN 6.
 OS avian infectious bronchitis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11120;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D1466;
 RA Lee C.-W., Jackwood M.W.;
 RT "Genetic evidence of recombination in the DE072 strain of coronavirus IBV.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF203006; AAF69123.1; -.
 DR InterPro; IPR001218; -.
 DR Pfam; PF00937; Corona_nucleoca; 1.
 SQ SEQUENCE 409 AA; 45176 MW; D85C3E379D2C6D40 CRC64;

Query Match 49.4%; Score 43; DB 14; Length 409;
 Best Local Similarity 53.3%; Pred. No. 24;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 PHRRGRPNRSRSKRT 16
 | : |||| : ||
 Db 336 PRPKSRPNRSRPATRT 350

Search completed: November 21, 2001, 16:27:39
 Job time: 403 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:20:49 ; Search time 44.87 Seconds
(without alignments)
18.915 Million cell updates/sec

Title: US-09-443-986A-3
Perfect score: 89
Sequence: 1 GTSNGGCCNYDGP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*

- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	89	100.0	15	AA03867	GIT receptor targete
2	49	55.1	488	AA03867	Arabidopsis thalia
3	49	55.1	503	AA03867	Arabidopsis thalia
4	48	53.9	490	AA03867	Arabidopsis thalia
5	43	48.3	127	AA03867	Human prostate tum
6	42	47.2	121	AA03867	Arabidopsis thalia
7	42	47.2	124	AA03867	Arabidopsis thalia
8	42	47.2	143	AA03867	zea mays protein f
9	42	47.2	158	AA03867	Arabidopsis thalia
10	41	46.1	194	AA03867	Human blood myocer
11	41	46.1	195	AA03867	Human ventricular

12	41	46.1	195	10	AA03867	Recombinant human
13	41	46.1	197	10	AA03867	Recombinant human
14	41	46.1	197	10	AA03867	Recombinant human
15	40.5	45.5	899	21	AA03867	Human mitogenic re
16	40	44.9	106	20	AA03867	Poxvirus ORF-2 tru
17	40	44.9	110	21	AA03867	Human epidermal pr
18	40	44.9	111	20	AA03867	Poxvirus ORF-2 tru
19	40	44.9	116	20	AA03867	Poxvirus ORF-2 tru
20	40	44.9	121	20	AA03867	Poxvirus ORF-2 tru
21	40	44.9	121	20	AA03867	Poxvirus ORF-2 (VE
22	40	44.9	148	20	AA03867	Poxvirus ORF-2 pro
23	40	44.9	148	20	AA03867	Parapox virus VEGF
24	39	43.8	188	20	AA03867	Arabidopsis thalia
25	39	43.8	107	21	AA03867	Arabidopsis thalia
26	39	43.8	109	21	AA03867	Arabidopsis thalia
27	39	43.8	112	21	AA03867	Arabidopsis thalia
28	39	43.8	114	21	AA03867	Arabidopsis thalia
29	39	43.8	116	21	AA03867	Arabidopsis thalia
30	39	43.8	118	21	AA03867	Arabidopsis thalia
31	39	43.8	165	21	AA03867	Arabidopsis thalia
32	39	43.8	165	21	AA03867	Arabidopsis thalia
33	39	43.8	165	21	AA03867	Arabidopsis thalia
34	39	43.8	207	21	AA03867	Arabidopsis thalia
35	39	43.8	211	21	AA03867	Arabidopsis thalia
36	39	43.8	234	21	AA03867	Arabidopsis thalia
37	39	43.8	301	21	AA03867	Arabidopsis thalia
38	39	43.8	301	21	AA03867	Arabidopsis thalia
39	39	43.8	367	21	AA03867	Arabidopsis thalia
40	39	43.8	367	21	AA03867	Arabidopsis thalia
41	39	43.8	373	21	AA03867	Arabidopsis thalia
42	39	43.8	374	21	AA03867	Arabidopsis thalia
43	39	43.8	378	21	AA03867	Human ORF2770
44	39	43.8	378	22	AA03867	Human GTP-binding
45	39	43.8	2799	19	AA03867	Human tumour suppr

ALIGNMENTS

RESULT 1	
AA03867	
ID	AA03867 standard; peptide; 15 AA.
XX	
AC	AA03867;
XX	
DT	20-OCT-2000 (first entry)
XX	
DE	GIT receptor targeting peptide zelan146 (HAX42 fragment).
XX	
KW	Retro-inversion peptide; gastrointestinal tract; active agent transport;
KW	GIT; hypertension; diabetes; osteoporosis; haemophilia; anaemia; cancer;
KW	migraine; angina pectoris.
XX	
OS	Synthetic.
XX	
PH	Key
FT	Modified-site 1
FT	Location/Qualifiers
FT	/note= "Lysine is dansylated"
FT	Misc-difference 2..14
FT	/note= "D-form residue"
XX	
PN	WO200031123-A2.
XX	
PD	02-JUN-2000.
XX	
PF	19-NOV-1999; 99WO-IE00117.
XX	
PR	19-NOV-1998; 98US-0109038.
XX	
PA	(ELAN-) ELAN CORP PLC.
XX	
PI	O'Mahony DJ;
XX	

DR WPI; 2000-400037/34.
 XX Retro-inverted peptide used to deliver active agents across the
 PT gastrointestinal tract to treat hypertension, diabetes, osteoporosis,
 PT haemophilia, anaemia, cancer, migraines and angina pectoris -
 XX
 FS Disclosure; Page 18; 36pp; English.
 XX
 CC This invention relates to retro-inverted peptides which specifically bind
 CC to the gastro-intestinal tract receptor HPn1, hPnPT1, D2H or hS1. Also
 CC included in the invention are a retro-inverted peptide which enhances the
 CC delivery of an active agent across the gastrointestinal tract (GI) into
 CC the systemic, portal or hepatic circulation. A composition comprising a
 CC retro-inverted peptide bound to a material comprising an active agent
 CC used to treat a mammalian disease or disorder is also disclosed in the
 CC invention. The retro-inversion peptides target gastrointestinal tract
 CC transport receptors to promote in vivo uptake of active agents and/or
 CC enhance active agent delivery across the tract into the systemic
 CC circulation. The gastrointestinal agents (containing retro-inverted
 CC peptides) are used to facilitate the transport of active ingredients
 CC through human or animal gastrointestinal tissue, from the lumen to the
 CC portal, hepatic, or systemic circulation. The compositions containing
 CC these agents can be used to treat or prevent mammalian, especially human,
 CC diseases or disorders, especially hypertension, diabetes, osteoporosis,
 CC haemophilia, anaemia, cancer, migraine, and angina pectoris. The
 CC compositions can be administered in vivo to image selected sites or
 CC tissues, such as the gastrointestinal tract, by using an imaging agent as
 CC the active agent.
 CC The present sequence represents a retro-inversion used in the invention.
 CC The sequence is a HAX42 14 mer fragment D form retro-inversion peptide.
 XX
 SQ Sequence 15 AA:

Query Match 100.0%; Score 89; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.le-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GTSNGNGCCNYDGP 14
 Db 2 gtsngngccnydgp 15
 |||||

RESULT 2
 AAG17315
 ID AAG17315 standard; Protein; 488 AA.
 AC AAG17315;
 XX
 XX 17-OCT-2000 (first entry)
 XX
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 18287.
 DE
 DE
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 XX Arabidopsis thaliana.
 XX
 XX EP1033405-A2.
 XX
 XX 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999; 99US-0121825.
 XX 05-MAR-1999; 99US-0123180.
 XX 09-MAR-1999; 99US-0123548.
 XX 23-MAR-1999; 99US-0125788.
 XX 25-MAR-1999; 99US-0126264.
 XX 29-MAR-1999; 99US-0126785.
 XX 01-APR-1999; 99US-0127462.
 XX 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.

PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 55.1%; Score 49; DB 21; Length 488;
Best Local Similarity 70.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 GNGCCNYDGP 14
| |||| ||
Db 16 gkgcckygpp 25

RESULT 3

AAG17314
ID AAG17314 standard; Protein; 503 AA.

XX AAG17314;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 18286.

DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 01-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142134.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148585.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0159584.

PR	21-OCT-1999;	99US-0160741.	PR	11-MAY-1999;	99US-0134256.
PR	21-OCT-1999;	99US-0160767.	PR	14-MAY-1999;	99US-0134218.
PR	21-OCT-1999;	99US-0160768.	PR	14-MAY-1999;	99US-0134219.
PR	21-OCT-1999;	99US-0160770.	PR	14-MAY-1999;	99US-0134221.
PR	21-OCT-1999;	99US-0160814.	PR	14-MAY-1999;	99US-0134370.
PR	21-OCT-1999;	99US-0160815.	PR	18-MAY-1999;	99US-0134768.
PR	22-OCT-1999;	99US-0160980.	PR	19-MAY-1999;	99US-0134941.
PR	22-OCT-1999;	99US-0160981.	PR	20-MAY-1999;	99US-0135124.
PR	22-OCT-1999;	99US-0160989.	PR	21-MAY-1999;	99US-0135353.
PR	22-OCT-1999;	99US-0160989.	PR	24-MAY-1999;	99US-0135629.
PR	25-OCT-1999;	99US-0161404.	PR	25-MAY-1999;	99US-0136021.
PR	25-OCT-1999;	99US-0161405.	PR	27-MAY-1999;	99US-0136392.
PR	26-OCT-1999;	99US-0161406.	PR	28-MAY-1999;	99US-0136782.
PR	26-OCT-1999;	99US-0161359.	PR	01-JUN-1999;	99US-0137222.
PR	26-OCT-1999;	99US-0161360.	PR	03-JUN-1999;	99US-0137528.
PR	26-OCT-1999;	99US-0161361.	PR	04-JUN-1999;	99US-0137502.
PR	28-OCT-1999;	99US-0161920.	PR	07-JUN-1999;	99US-0137724.
PR	28-OCT-1999;	99US-0161992.	PR	08-JUN-1999;	99US-0138094.
PR	28-OCT-1999;	99US-0161993.	PR	10-JUN-1999;	99US-0138540.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;	99US-0138847.
Query Match 55.1%; Score 49; DB 21; Length 503;					
Best Local Similarity 70.0%; Pred. No. 46;					
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;					
QY	5	GNCCNYDGP 14			
Db	31	gkcckyggp 40			
RESULT 4					
AAAG36718	AAAG36718 standard; Protein; 490 AA.				
XX	AC	AAAG36718;			
XX	CC	{			
DT	18-OCT-2000	(first entry)			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 45038.				
XX	Protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
XX	Arabidopsis thaliana.				
OS	Arabidopsis thaliana.				
XX	EP1033405-A2.				
PN	06-SEP-2000.				
PD	25-FEB-2000; 2000EP-0301439.				
XX	25-FEB-1999;	99US-0121825.			
PR	05-MAR-1999;	99US-0123180.			
PR	09-MAR-1999;	99US-0123548.			
PR	23-MAR-1999;	99US-0125788.			
PR	25-MAR-1999;	99US-0126264.			
PR	29-MAR-1999;	99US-0126785.			
PR	01-APR-1999;	99US-0127462.			
PR	06-APR-1999;	99US-0128234.			
PR	08-APR-1999;	99US-0128714.			
PR	16-APR-1999;	99US-0129845.			
PR	19-APR-1999;	99US-0130077.			
PR	21-APR-1999;	99US-0130449.			
PR	23-APR-1999;	99US-0130510.			
PR	28-APR-1999;	99US-0130891.			
PR	30-APR-1999;	99US-0131449.			
PR	30-APR-1999;	99US-0132048.			
PR	04-MAY-1999;	99US-0132407.			
PR	04-MAY-1999;	99US-0132484.			
PR	05-MAY-1999;	99US-0132485.			
PR	06-MAY-1999;	99US-0132486.			
PR	06-MAY-1999;	99US-0132487.			
PR	07-MAY-1999;	99US-0132863.			
PR	11-MAY-1999;	99US-0139119.			
PR	14-MAY-1999;	99US-0139452.			
PR	16-JUN-1999;	99US-0139453.			
PR	17-JUN-1999;	99US-0139452.			
PR	18-JUN-1999;	99US-0139454.			
PR	18-JUN-1999;	99US-0139455.			
PR	18-JUN-1999;	99US-0139456.			
PR	18-JUN-1999;	99US-0139457.			
PR	18-JUN-1999;	99US-0139458.			
PR	18-JUN-1999;	99US-0139459.			
PR	18-JUN-1999;	99US-0139460.			
PR	18-JUN-1999;	99US-0139461.			
PR	18-JUN-1999;	99US-0139462.			
PR	18-JUN-1999;	99US-0139463.			
PR	18-JUN-1999;	99US-0139750.			
PR	18-JUN-1999;	99US-0139763.			
PR	21-JUN-1999;	99US-0139817.			
PR	22-JUN-1999;	99US-0139899.			
PR	23-JUN-1999;	99US-0140353.			
PR	23-JUN-1999;	99US-0140354.			
PR	24-JUN-1999;	99US-0140695.			
PR	28-JUN-1999;	99US-0140823.			
PR	29-JUN-1999;	99US-0140991.			
PR	30-JUN-1999;	99US-0141287.			
PR	01-JUL-1999;	99US-0141842.			
PR	01-JUL-1999;	99US-0142154.			
PR	02-JUL-1999;	99US-0142055.			
PR	06-JUL-1999;	99US-0142390.			
PR	08-JUL-1999;	99US-0142803.			
PR	09-JUL-1999;	99US-0142920.			
PR	12-JUL-1999;	99US-0142977.			
PR	13-JUL-1999;	99US-0143542.			
PR	14-JUL-1999;	99US-0143624.			
PR	15-JUL-1999;	99US-0144005.			
PR	16-JUL-1999;	99US-0144085.			
PR	16-JUL-1999;	99US-0144086.			
PR	19-JUL-1999;	99US-0144325.			
PR	19-JUL-1999;	99US-0144331.			
PR	19-JUL-1999;	99US-0144332.			
PR	19-JUL-1999;	99US-0144333.			
PR	19-JUL-1999;	99US-0144334.			
PR	19-JUL-1999;	99US-0144335.			
PR	20-JUL-1999;	99US-0144352.			
PR	20-JUL-1999;	99US-0144632.			
PR	20-JUL-1999;	99US-0144884.			
PR	21-JUL-1999;	99US-0144814.			
PR	21-JUL-1999;	99US-0145086.			
PR	21-JUL-1999;	99US-0145088.			
PR	22-JUL-1999;	99US-0145085.			
PR	22-JUL-1999;	99US-0145087.			
PR	22-JUL-1999;	99US-0145089.			
PR	22-JUL-1999;	99US-0145192.			
PR	23-JUL-1999;	99US-0145145.			

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.9%; Score 48; DB 21; Length 490;
Best Local Similarity 64.3%; Pred. No. 61;
Matches 9; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 3 SNG--NGCCNYDGP 14
Db 14 snggskgckkyggp 27

RESULT 5

AAV73985
ID AAV73985 standard; Protein; 127 AA.

XX AC AAV73985;

XX DT 14-MAR-2000 (first entry)

XX DE Human prostate tumor EST fragment derived protein #172.

XX KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
XX KW treatment.

XX OS Homo sapiens.

XX PN DE19820190-A1.

XX PD 04-NOV-1999.

XX PF 28-APR-1998; 98DE-1020190.

XX PR 28-APR-1998; 98DE-1020190.

XX PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

XX DR WPI; 1999-621386/54.

XX DR N-PSDB; AAZ52914.

XX PT New human nucleic acid sequences from pancreatic tumors, and related
XX PT proteins -

XX PS Claim 23; Page 381; 502pp; German.

XX CC This invention describes novel polypeptides and their encoding nucleic
XX CC acids derived from human pancreatic tumor tissue which have cytostatic
XX CC activity. The sequences are also useful in producing pharmaceutical
XX CC compositions for treatment of pancreatic tumors. AAV73814-Y74252
XX CC represent protein fragments encoded by the human pancreatic tumor cDNA
XX CC library derived expressed sequence tag (EST) sequences represented in
XX CC AAZ52858-Z53014.

XX SQ Sequence 127 AA;

Query Match 48.3%; Score 43; DB 20; Length 127;
Best Local Similarity 50.0%; Pred. No. 82;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYDGP 14


```
Db      18 gtagstcrdypgp 31
RESULT  6
AAG59067
ID      AAG59067 standard; Protein; 121 AA.
AC      AAG59067;
XX
DT      18-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana protein fragment SEQ ID NO: 76362.
XX
KW      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KW      termination sequence.
XX
OS      Arabidopsis thaliana.
XX
PN      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-0301439.
XX
PR      25-FEB-1999; 99US-0121825.
PR      05-MAR-1999; 99US-0123180.
PR      09-MAR-1999; 99US-0123548.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127452.
PR      06-APR-1999; 99US-0128234.
PR      08-APR-1999; 99US-0128714.
PR      16-APR-1999; 99US-0129845.
PR      19-APR-1999; 99US-0130077.
PR      21-APR-1999; 99US-0130449.
PR      23-APR-1999; 99US-0130510.
PR      23-APR-1999; 99US-0130891.
PR      28-APR-1999; 99US-0131449.
PR      30-APR-1999; 99US-0132048.
PR      30-APR-1999; 99US-0132407.
PR      04-MAY-1999; 99US-0132484.
PR      05-MAY-1999; 99US-0132485.
PR      06-MAY-1999; 99US-0132486.
PR      06-MAY-1999; 99US-0132487.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134221.
PR      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139494.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      18-JUN-1999; 99US-0139817.
PR      21-JUN-1999; 99US-0139899.
PR      22-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      23-JUN-1999; 99US-0140695.
PR      24-JUN-1999; 99US-0140823.
PR      28-JUN-1999; 99US-0140991.
PR      29-JUN-1999; 99US-0141287.
PR      30-JUN-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      06-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
PR      09-AUG-1999; 99US-0147493.
PR      09-AUG-1999; 99US-0147935.
PR      10-AUG-1999; 99US-0148171.
PR      11-AUG-1999; 99US-0148319.
PR      12-AUG-1999; 99US-0148341.
PR      13-AUG-1999; 99US-0148565.
PR      13-AUG-1999; 99US-0148684.
PR      16-AUG-1999; 99US-0149368.
```


PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.2%; Score 42; DB 21; Length 124;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 CCNYDGP 14
||| |||
Db 101 ccnvdp 107

RESULT 8
AAG44635
ID AAG44635 standard; Protein; 143 AA.
XX AC AAG44635;
XX DT 18-OCT-2000 (first entry)
XX DE Zea mays protein fragment SEQ ID NO: 55935.
XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.

OS Zea mays subsp. mays.

XX EP1033405-A2.

PN

XX

PD

XX

PF

XX

PR 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126788.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR	01-SEP-1999;	99US-0151930.	XX	06-SEP-2000.	
PR	07-SEP-1999;	99US-0152363.	PD		
PR	10-SEP-1999;	99US-0153070.	XX		
PR	13-SEP-1999;	99US-0153758.	PF	25-FEB-2000; 2000EP-0301439.	
PR	15-SEP-1999;	99US-0154018.	XX		
PR	16-SEP-1999;	99US-0154039.	PR	25-FEB-1999;	99US-0121825.
PR	20-SEP-1999;	99US-0154779.	PR	05-MAR-1999;	99US-0123180.
PR	22-SEP-1999;	99US-0155139.	PR	09-MAR-1999;	99US-0123548.
PR	23-SEP-1999;	99US-0155486.	PR	23-MAR-1999;	99US-0125788.
PR	24-SEP-1999;	99US-0155659.	PR	25-MAR-1999;	99US-0126264.
PR	28-SEP-1999;	99US-0156458.	PR	29-MAR-1999;	99US-0126785.
PR	29-SEP-1999;	99US-0156596.	PR	01-APR-1999;	99US-0127462.
PR	04-OCT-1999;	99US-0157117.	PR	06-APR-1999;	99US-0128234.
PR	05-OCT-1999;	99US-0157753.	PR	08-APR-1999;	99US-0128714.
PR	06-OCT-1999;	99US-0157865.	PR	16-APR-1999;	99US-0129845.
PR	07-OCT-1999;	99US-0158029.	PR	19-APR-1999;	99US-0130077.
PR	08-OCT-1999;	99US-0158232.	PR	21-APR-1999;	99US-0130449.
PR	12-OCT-1999;	99US-0158369.	PR	23-APR-1999;	99US-0130510.
PR	13-OCT-1999;	99US-0159293.	PR	28-APR-1999;	99US-0130891.
PR	13-OCT-1999;	99US-0159294.	PR	30-APR-1999;	99US-0131449.
PR	13-OCT-1999;	99US-0159295.	PR	30-APR-1999;	99US-0132048.
PR	14-OCT-1999;	99US-0159329.	PR	04-MAY-1999;	99US-0132407.
PR	14-OCT-1999;	99US-0159330.	PR	06-MAY-1999;	99US-0132484.
PR	14-OCT-1999;	99US-0159331.	PR	06-MAY-1999;	99US-0132485.
PR	14-OCT-1999;	99US-0159637.	PR	07-MAY-1999;	99US-0132486.
PR	14-OCT-1999;	99US-0159638.	PR	11-MAY-1999;	99US-0132863.
PR	18-OCT-1999;	99US-0159584.	PR	14-MAY-1999;	99US-0134218.
PR	21-OCT-1999;	99US-0160741.	PR	14-MAY-1999;	99US-0134219.
PR	21-OCT-1999;	99US-0160767.	PR	14-MAY-1999;	99US-0134221.
PR	21-OCT-1999;	99US-0160768.	PR	14-MAY-1999;	99US-0134370.
PR	21-OCT-1999;	99US-0160770.	PR	18-MAY-1999;	99US-0134768.
PR	21-OCT-1999;	99US-0160814.	PR	19-MAY-1999;	99US-0134941.
PR	21-OCT-1999;	99US-0160815.	PR	20-MAY-1999;	99US-0135124.
PR	22-OCT-1999;	99US-0160980.	PR	21-MAY-1999;	99US-0135353.
PR	22-OCT-1999;	99US-0160981.	PR	24-MAY-1999;	99US-0135629.
PR	22-OCT-1999;	99US-0160989.	PR	25-MAY-1999;	99US-0136021.
PR	25-OCT-1999;	99US-0161404.	PR	27-MAY-1999;	99US-0136392.
PR	25-OCT-1999;	99US-0161405.	PR	28-MAY-1999;	99US-0136782.
PR	25-OCT-1999;	99US-0161406.	PR	01-JUN-1999;	99US-0137222.
PR	26-OCT-1999;	99US-0161359.	PR	03-JUN-1999;	99US-0137528.
PR	26-OCT-1999;	99US-0161360.	PR	04-JUN-1999;	99US-0137502.
PR	26-OCT-1999;	99US-0161361.	PR	07-JUN-1999;	99US-0137724.
PR	28-OCT-1999;	99US-0161920.	PR	08-JUN-1999;	99US-0138094.
PR	28-OCT-1999;	99US-0161992.	PR	10-JUN-1999;	99US-0138540.
PR	28-OCT-1999;	99US-0161993.	PR	14-JUN-1999;	99US-0138847.
PR	29-OCT-1999;	99US-0162142.	PR	16-JUN-1999;	99US-0139119.
Query Match 47.2%; Score 42; DB 21; Length 143;					
Best Local Similarity 54.5%; Pred. No. 1.2e+02;					
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;					
Qy	4	NGNGCCNVDPG 14			
		:			
Db	99	nhshchhyggp 109			
RESULT 9					
AGS9065					
ID	AGS9065	standard; Protein; 158 AA.			
XX					
AC	AGS9065;				
XX					
DT	18-OCT-2000	(first entry)			
XX					
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 76360.			
XX					
KW	Protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
XX					
OS	Arabidopsis thaliana.				
XX					
PN	EPI033405-A2.				

PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-01451145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.

PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.2%; Score 42; DB 21; Length 158;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 CCNVDGP 14
||| |||
Db 135 ccnvdgp 141

RESULT 10
AAY49916
ID AAY49916 standard; Protein; 194 AA.

XX AC AAY49916;

XX DT 27-JAN-2000 (first entry)

DE Human blood myocardial myosin light chain I.

KW Human; blood; cardiac muscle myosin light chain I; diagnosis;
KW myocardial myosin light chain I; acute myocardial infarction;
KW antibody; antigen.

OS Homo sapiens.

XX CN125839-A.

XX 18-AUG-1999.

XX 04-DEC-1998; 98CN-0122066.

XX 04-DEC-1998; 98CN-0122066.

XX

PA (SHAN-) SHANGHAI BIO-CHEM INST CHINESE ACAD SCI.

PI Gong Z, Peng B, Zhou G;

XX WPI; 1999-591529/51.

DR N-PSDB; AA235675.

XX Diagnosis reagent for blood cardiac muscle myosin light chain I - used
PT in a double-antibody sandwich method

XX Claim 3; Page 1; 18pp; Chinese.

XX The present sequence represents human blood myocardial myosin light
CC chain I. The blood myocardial myosin light chain I diagnostic reagent
CC mainly includes the high expression product of human myocardial myosin
CC light-chain I gene in the colibacillus as positive control and single
CC antibody and multiple antibody prepared by using the expression product
CC as antigen. The diagnostic method is a double-antibody sandwich
CC method, i.e. it uses the immobilised single antibody to trap the
CC antigen being in the tested serum-myocardial myosin light-chain I, and
CC uses the multiple antibody as testing antibody, so that according to
CC the ELISA reading value measured after the reaction of enzyme and
CC substrate, and cut off value provided by the invention, if the value is
CC greater than cut off value, it is determined as pathogenic stage of
CC acute myocardial infarction.

XX Sequence 194 AA;

Query Match 46.1%; Score 41; DB 20; Length 194;

Best Local Similarity 50.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12

DB 175 gqedsngcinye 186

RESULT 11

AAAP91391

ID AAP91391 standard; protein; 195 AA.

XX AC

XX AC

XX DT 08-FEB-1990 (first entry)

XX Human ventricular myosin light chain 1.

XX Cardiac myosin; myocardial infarction; light chain 1.

XX OS Homo sapiens.

XX PN EP336155-A.

XX PD 11-OCT-1989.

XX PF 15-MAR-1989; 89EP-0104607.

XX PR 17-MAR-1988; 88JP-0064142.

XX PA (YAMS) YAMASA SHOYO KK.

XX PI Kurabayashi M, Yazaki Y, Noguchi T, Ohtani Y, Kato H, Sugl M;

XX WPI; 1989-294079/41.

XX DR N-PSDB; AA235675.

XX PT Recombinant cardiac myosin protein of human origin - is used as antigen

XX PT to obtain antibodies or as standard to measure cardiac myosin concn. in

XX PT assays.

XX PS Disclosure; fig 2B; 36pp; English.

CC Sequence codes for atrial myosin light chain I. The protein is used in
CC assays for myocardial infarction.

SQ Sequence 195 AA;

Query Match 46.1%; Score 41; DB 10; Length 195;

Best Local Similarity 50.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12

DB 175 gqedsngcinye 186

RESULT 12

AAAP90745

ID AAP90745 standard; protein; 195 AA.

XX AC

XX AC

XX DT 12-FEB-1990 (first entry)

XX DE Recombinant human cardiac myosin light chain 1.

XX KW Cardiac myosin; myocardial infarction; light chain 1.

XX OS Homo sapiens.

XX FH Key

XX FT region 11

XX FT region 12

XX FT region 15

XX FT region 18

XX FT region 19

XX PN EP336155-A.

XX PD 11-OCT-1989.

XX PF 15-MAR-1989; 89EP-0104607.

XX PR 17-MAR-1988; 88JP-0064142.

XX PA (YAMS) YAMASA SHOYO KK.

XX PI Kurabayashi M, Yazaki Y, Noguchi T, Ohtani Y, Kato H, Sugl M;

XX WPI; 1989-294079/41.

XX PT Recombinant cardiac myosin protein of human origin - is used as antigen

XX PT to obtain antibodies or as standard to measure cardiac myosin concn. in

XX PT assays.

XX PS Claim 5; page 21; 36pp; English.

XX CC Sequence contains code for cardiac myosin. The protein is used in

XX CC assays for myocardial infarction. AA11= Glu or Asp; AA12= Ala or Asp;

XX CC AA15= Pro or Ala; AA18= Ala or Lys; and AA19= Pro or Ala.

XX CC See also AAP90746.

XX SQ Sequence 195 AA;

Query Match 46.1%; Score 41; DB 10; Length 195;

Best Local Similarity 50.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12

DB 175 gqedsngcinye 186

```

RESULT 13
AAP91390
ID AAP91390 standard; protein; 197 AA.
XX
AC AAP91390;
XX
DT 12-FEB-1990 (first entry)
XX
DE Recombinant human cardiac myosin light chain 1.
XX
KW Cardiac myosin; myocardial infarction; light chain 1.
XX
OS Homo sapiens.
XX
PN EP336155-A.
XX
PD 11-OCT-1989.
XX
PF 15-MAR-1989; 89EP-0104607.
XX
PR 17-MAR-1988; 88JP-0064142.
XX
PA (YAMS ) YAMASA SHOYO KK.
XX
PI Kurabayashi M, Yazaki Y, Noguchi T, Ohtani Y, Kato H, Sugi M;
XX
WPI; 1989-294079/41.
XX
Recombinant cardiac myosin protein of human origin - is used as antigen
to obtain antibodies or as standard to measure cardiac myosin concn. in
assays.
XX
PS Claim 5; page 21; 36pp; English.
XX
CC Sequence contains code for cardiac myosin. The protein is used in
assays for myocardial infarction. AA11= Glu or Asp; AA12= Ala or Asp;
AA15= Pro or Ala; AA18= Ala or Lys; and AA19= Pro or Ala.
XX
CC See also AAP90745.
XX
SQ Sequence 197 AA;

Query Match 46.1%; Score 41; DB 10; Length 197;
Best Local Similarity 50.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12
| : ||| ||:
Db 177 gqedangcinye 188

RESULT 15
AAAY71124
ID AAAY71124 standard; Protein; 899 AA.
XX
AC AAAY71124;
XX
XX 08-SEP-2000 (first entry)
XX
DE Human mitogenic regulator duox2.
XX
KW Human; mitogenic regulator; duox2; dual oxidase; hydrogen peroxide;
KW superoxide; reactive oxygen intermediate; ROI; cell division;
KW cytosolic; antipsoriatic; cardiant; antiarteriosclerotic; vasotropic;
KW antiangiogenic; hypotensive; drug development; treatment; cancer;
KW abnormal growth; psoriasis; prostatic hypertrophy; proliferation;
KW blood vessel; lymphatic vessel; benign prostatic hypertrophy;
KW cardiovascular disease; arteriovenous malformation; eye disorder;
KW hypertension; atherosclerosis; restenosis; angioplasty.
XX
OS Homo sapiens.
XX
PN WO200028031-A2.
XX
PD 18-MAY-2000.
XX
PF 10-NOV-1999; 99WO-US26592.
XX
PR 10-NOV-1998; 98US-0107911.
XX
PR 17-AUG-1999; 99US-0149332.
XX
PR 27-AUG-1999; 99US-0151242.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Lambeth JD, Lassegue BP, Griendling KK, Arnold RS, Guangjie C;
XX
WPI; 2000-376545/32.
XX
N-PSDB; AAD00697.

```


XX Protein capable of stimulating superoxide production, useful for
PT treating conditions associated with abnormal growth, including cancer
PT
XX
PS Claim 2; Page 134-138; 141pp; English.
XX
CC The present sequence is a human dual oxidase duox2,
CC which is capable of catalysing both the generation of superoxide and
CC peroxidative type reactions. The superoxide and hydrogen peroxide are
CC reactive oxygen intermediates (ROI) that affect cell division.
CC The duox2 protein functions as a mitogenic regulator.
CC It has three distinct regions, the N-terminal
CC region having homology to peroxidase proteins, the internal region
CC having homology to calmodulin proteins (CAM) and the C-terminal region
CC having homology to mox proteins. Duox2 shows highest expression in adult
CC lung and foetal tissues. The present sequence is useful in developing
CC drugs and therapies for treatment of conditions associated with abnormal
CC growth, including cancer, psoriasis, prostatic hypertrophy, benign
CC prostatic hypertrophy, cardiovascular disease, proliferation of vessels,
CC e.g. blood vessels and lymphatic vessels, arteriovenous malformation,
CC vascular problems associated with eye, atherosclerosis, hypertension,
CC and restenosis following angioplasty.
XX
SQ Sequence 899 AA;

Query Match 45.5%; Score 40.5; DB 21; Length 899;
Best Local Similarity 53.3%; Pred. No. 1e+03;
Matches 8; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY 5 GNGCCNY-----DGP 14
|||||
Db 703 gngcagypklyldgp 717

Search completed: November 21, 2001, 16:20:50
Job time: 300 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:21:21 : Search time 24.83 Seconds
(without alignments)
12,688 Million cell updates/sec

Title: US-09-443-986A-3
Perfect score: 89
Sequence: 1 GTSNGNGCCNYDGP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2.6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2.6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2.6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2.6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2.6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	47.2	730	4	US-08-872-757-2
2	40	44.9	58	4	US-08-963-851-25
3	39	43.8	20	4	US-09-314-242-10
4	38.5	43.3	280	2	US-08-284-941-7
5	38.5	43.3	280	2	US-08-447-642-7
6	38.5	43.3	280	4	US-09-236-503-7
7	38.5	43.3	280	5	PCT-US93-02147A-7
8	38.5	43.3	355	2	US-08-865-203-5
9	38.5	43.3	355	2	US-07-849-420-5
10	38.5	43.3	355	4	US-09-253-854-5
11	38.5	43.3	355	4	US-08-955-424-5
12	38.5	43.3	621	1	US-08-328-961-2
13	38.5	43.3	621	1	US-08-462-397-2
14	38.5	43.3	783	2	US-08-545-562A-5
15	38	42.7	109	4	US-09-124-671-23
16	38	42.7	451	1	US-08-191-337-3
17	38	42.7	452	1	US-08-290-978A-5
18	38	42.7	452	2	US-08-780-869-5
19	37	41.6	15	4	US-08-481-968A-17
20	37	41.6	22	1	US-08-423-399B-27
21	37	41.6	24	4	US-09-443-501A-22
22	37	41.6	447	1	US-08-365-689-1
23	37	41.6	447	1	US-07-747-781-1
24	37	41.6	447	1	US-08-145-138A-1
25	37	41.6	447	5	PCT-US92-06888-1
26	37	41.6	447	5	PCT-US93-03027-5
27	37	41.6	492	4	US-08-462-794-11

28	37	41.6	899	1	US-08-365-689-2	Sequence 2, Appl
29	37	41.6	899	1	US-08-145-138A-2	Sequence 2, Appl
30	37	41.6	933	1	US-07-747-781-2	Sequence 2, Appl
31	37	41.6	933	5	PCT-US92-06888-2	Sequence 2, Appl
32	37	41.6	972	3	US-08-750-141A-2	Sequence 2, Appl
33	37	41.6	979	1	US-08-308-881-6	Sequence 6, Appl
34	37	41.6	979	2	US-09-058-263-6	Sequence 6, Appl
35	37	41.6	979	2	US-09-059-099-6	Sequence 6, Appl
36	37	41.6	979	3	US-09-058-264-6	Sequence 6, Appl
37	37	41.6	979	5	PCT-US95-06530-6	Sequence 6, Appl
38	37	41.6	1345	2	US-08-377-767-3	Sequence 3, Appl
39	37	41.6	2368	1	US-08-198-446B-15	Sequence 15, Appl
40	37	41.6	2368	2	US-08-870-693-15	Sequence 15, Appl
41	37	41.6	4654	4	US-08-476-515A-84	Sequence 84, Appl
42	37	41.6	4655	4	US-08-652-877-84	Sequence 84, Appl
43	37	41.6	4655	4	US-08-852-877-86	Sequence 86, Appl
44	37	41.6	4655	4	US-08-852-877-88	Sequence 88, Appl
45	37	41.6	4655	4	US-08-652-877-90	Sequence 90, Appl

ALIGNMENTS

RESULT 1
US-08-872-757-2
; Sequence 2, Application US/08872757
; Patent No. 6258584
; GENERAL INFORMATION:
; APPLICANT: Prockop, Darwin J.
; APPLICANT: Hojima, Yoshio
; APPLICANT: Li, Shi-Wu
; APPLICANT: Sieron, Aleksander
; TITLE OF INVENTION: RECOMBINANT C-PROTEINASE AND
; TITLE OF INVENTION: PROCESSES; METHODS AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,757
; FILING DATE: 10-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,187
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 8389-028-999
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-872-757-2

Query Match 47.2%; Score 42; DB 4; Length 730;
Best Local Similarity 75.0%; Pred. No. 2.7e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 GCGNYDGP 14

|||||

Db 184 GCCSYVGP 191

RESULT 2

US-08-963-851-25

; Sequence 25, Application US/08963851

; Patent No. 6300116

; GENERAL INFORMATION:

; APPLICANT: VAN DER OSTEN, CLAUS

; APPLICANT: HALKIER, TORDEN

; APPLICANT: ANDERSEN, CARSTEN

; APPLICANT: BAUDITZ, PETER

; APPLICANT: HANSEN, PETER KAMP

; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS

; FILE REFERENCE: 4946,200-US

; CURRENT APPLICATION NUMBER: US/08/963,851

; CURRENT FILING DATE: 1997-11-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 25

; LENGTH: 58

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-08-963-851-25

Query Match 44.9%; Score 40; DB 4; Length 58;

Best Local Similarity 47.1%; Pred. No. 48;

Matches 8; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

Qy 1 GTSNGNC-----CNYDG 13

|||

Db 23 GVTEGGTGRGDCNYDG 39

RESULT 3

US-09-314-242-10

; Sequence 10, Application US/09314242A

; Patent No. 6248575

; GENERAL INFORMATION:

; APPLICANT: Elizabeth J. Golightly

; TITLE OF INVENTION: Nucleic Acids Encoding Polypeptides

; TITLE OF INVENTION: Having L-Amino Acid Oxidase Activity

; FILE REFERENCE: 5556,200-US

; CURRENT APPLICATION NUMBER: US/09/314,242A

; CURRENT FILING DATE: 1999-05-18

; EARLIER APPLICATION NUMBER: 09/080,428

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 10

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Trichoderma harzianum

US-09-314-242-10

Query Match 43.8%; Score 39; DB 4; Length 20;

Best Local Similarity 75.0%; Pred. No. 24;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTSNGNC 8

|||

Db 13 GTNGGNC 20

RESULT 4

US-08-284-941-7

; Sequence 7, Application US/08284941

; Patent No. 5863756

; GENERAL INFORMATION:

; APPLICANT: BARR, PHILIP J

; APPLICANT: KIEFER, MICHAEL C

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND

; TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM

; STREET: FIVE PALO ALTO SQUARE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/284,941

; FILING DATE: 2 August 1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: NEELEY PH.D., RICHARD L.

; REGISTRATION NUMBER: 30092

; REFERENCE/DOCKET NUMBER: CHIR-009/0105

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 843-5070

; TELEFAX: (415) 857-0663

; TELEX: 380816 COOLEY PA

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 280 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-284-941-7

Query Match 43.3%; Score 38.5; DB 2; Length 280;

Best Local Similarity 50.0%; Pred. No. 3.2e+02;

Matches 8; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

Qy 3 SNGNG-----CCNYDG 13

|||||

Db 178 ASGNGGTGRGDCNYDG 193

RESULT 5

US-08-447-642-7

; Sequence 7, Application US/08447642

; Patent No. 5989890

; GENERAL INFORMATION:

; APPLICANT: BARR, PHILIP J

; APPLICANT: KIEFER, MICHAEL C

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND

; TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM

; STREET: FIVE PALO ALTO SQUARE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/447,642

; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: US 08/284,941
; FILING DATE: 2 August 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY PH.D., RICHARD L.
; REGISTRATION NUMBER: 30092
; REFERENCE/DOCKET NUMBER: CHIR-009/01US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 843-5070
; TELEFAX: (415) 857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-447-642-7

Query Match 43.3%; Score 38.5; DB 2; Length 280;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

Qy 3 SNGNG-----CCNYDG 13
:|||||
Db 178 ASGNGGTRGDCNYDG 193

RESULT 6
US-09-236-503-7
; Sequence 7, Application US/09236503
; Patent No. 6277590
; GENERAL INFORMATION:
; APPLICANT: Barr, Philip J
; APPLICANT: Kiefer, Michael C
; TITLE OF INVENTION: Compositions and Methods for PACE 4 and 4.1 Gene and
; TITLE OF INVENTION: Polypeptides in Cells
; FILE REFERENCE: CHIR-009/04US
; CURRENT APPLICATION NUMBER: US/09/236,503
; CURRENT FILING DATE: 1999-01-25
; EARLIER APPLICATION NUMBER: 08/447,642
; EARLIER FILING DATE: 1995-05-23
; EARLIER APPLICATION NUMBER: 08/284,941
; EARLIER FILING DATE: 1994-08-02
; EARLIER APPLICATION NUMBER: 07/848,629
; EARLIER FILING DATE: 1992-03-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-236-503-7

Query Match 43.3%; Score 38.5; DB 4; Length 280;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

Qy 3 SNGNG-----CCNYDG 13
:|||||
Db 178 ASGNGGTRGDCNYDG 193

RESULT 7
PCT-US93-02147A-7
; Sequence 7, Application PC/TUS9302147A
; GENERAL INFORMATION:
; APPLICANT: BARR, PHILIP J
; APPLICANT: KIEFER, MICHAEL C

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND
; TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
; STREET: FIVE PALO ALTO SQUARE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/02147A
; FILING DATE: 19930309
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,629
; FILING DATE: 09-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY PH.D., RICHARD L
; REGISTRATION NUMBER: 30092
; REFERENCE/DOCKET NUMBER: CHIR-009/00US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-7622
; TELEFAX: (415) 857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-02147A-7

Query Match 43.3%; Score 38.5; DB 5; Length 280;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

Qy 3 SNGNG-----CCNYDG 13
:|||||
Db 178 ASGNGGTRGDCNYDG 193

RESULT 8
US-08-865-203-5
; Sequence 5, Application US/08865203
; Patent No. 5935815
; GENERAL INFORMATION:
; APPLICANT: van de Ven, Willem Jan Marie
; APPLICANT: van den Ouweland, Anna Maria Wilhelmina
; APPLICANT: Van Duijnhoven, Johannes Lambertus Petrus
; APPLICANT: Robroek, Antonius Johannes Maria
; APPLICANT: Konig, Piet Nico Maria
; TITLE OF INVENTION: Pharmaceutical Composition Having An
; TITLE OF INVENTION: Endoproteolytic Activity; A Process for
; TITLE OF INVENTION: Endoproteolytically Processing (Precursor)
; TITLE OF INVENTION: Proteins And For The (Micro)Biological
; TITLE OF INVENTION: Production Of Proteins
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOFFMANN & BARON, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,203
FILING DATE: 29-MAY-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Tran, Jessica H.
REGISTRATION NUMBER: 40,846
REFERENCE/DOCKET NUMBER: 294-41 DIV II
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-865-203-5

Query Match 43.3%; Score 38.5; DB 2; Length 355;
Best Local Similarity 50.0%; Pred. No. 3.9e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

QY 3 SNGNG-----CCNYDG 13
Db 184 ASNGGTRGDCNYDG 199

RESULT 9
US-07-849-420-5
Sequence 5, Application US/07849420
Patent No. 5989856
GENERAL INFORMATION:
APPLICANT: van de Ven, Willem Jan Marie
APPLICANT: van den Ouweland, Anna Maria Wilhelmina
APPLICANT: Van Duijnhooven, Johannes Lambertus Petrus
APPLICANT: Robroek, Antonius Johannes Maria
APPLICANT: Koning, Piet Nico Maria
TITLE OF INVENTION: Pharmaceutical Composition Having An
ENDOPROTEOLYTIC ACTIVITY; A PROCESS FOR
ENDOPROTEOLYTICALLY PROCESSING (PRECURSOR)
TITLE OF INVENTION: Endoproteolytically Processing (Precursor)
TITLE OF INVENTION: Proteins And For The (Micro)Biological
TITLE OF INVENTION: Production Of Proteins
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/849,420
FILING DATE: 19920624
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Moran, Thomas F.
REGISTRATION NUMBER: 16,579
REFERENCE/DOCKET NUMBER: 2805/41413
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9550

TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-849-420-5

Query Match 43.3%; Score 38.5; DB 2; Length 355;
Best Local Similarity 50.0%; Pred. No. 3.9e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

QY 3 SNGNG-----CCNYDG 13
Db 184 ASNGGTRGDCNYDG 199

RESULT 10
US-09-253-854-5
Sequence 5, Application US/09253854
Patent No. 6132717
GENERAL INFORMATION:
APPLICANT: van de Ven, Willem Jan Marie;
APPLICANT: van den Ouweland, Anna Maria Wilhelmina;
APPLICANT: Van Duijnhooven, Johannes Lambertus Petrus;
APPLICANT: Robroek, Antonius Johannes Maria; and
APPLICANT: Koning, Piet Nico Maria
TITLE OF INVENTION: Pharmaceutical Composition Having An
ENDOPROTEOLYTIC ACTIVITY; A PROCESS FOR
ENDOPROTEOLYTICALLY PROCESSING (PRECURSOR)
TITLE OF INVENTION: Endoproteolytically Processing (Precursor)
TITLE OF INVENTION: Proteins And For The (Micro)Biological
TITLE OF INVENTION: Production Of Proteins
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOFFMANN & BARON, LLP
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: U.S.A.
ZIP: 11753
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/253,854
FILING DATE: Unassigned
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Tran, Jessica H.
REGISTRATION NUMBER: 40,846
REFERENCE/DOCKET NUMBER: 294-41 DIV II/CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-253-854-5

Query Match 43.3%; Score 38.5; DB 4; Length 355;
Best Local Similarity 50.0%; Pred. No. 3.9e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

```
Qy 3 SNGNG-----CCNYDG 13
:|||||
Db 184 ASGNGGTRGDCNYDG 199

RESULT 11
US-08-955-424-5
: Sequence 5, Application US/08955424
: Patent No. 6274365
: GENERAL INFORMATION:
: APPLICANT: Van de Ven, Willem Jan Marie
: APPLICANT: Van de Ouweland, Anna Maria Wilhelmina
: APPLICANT: Van Duijnhoven, Johannes Lambertus Petrus
: APPLICANT: Koning, Piet Nico Maria
: APPLICANT: Roebroek, Antonius Johannes Maria
: TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION HAVING AN ENDOPEPTOLYTIC
: TITLE OF INVENTION: ACTIVITY; A PROCESS FOR ENDOPEPTOLYTICALLY PROCESSING
: TITLE OF INVENTION: (PRECURSOR) PROTEINS AND FOR THE (MICRO)BIOLOGICAL
: TITLE OF INVENTION: PRODUCTION OF PROTEINS
: FILE REFERENCE: SEQUENCE LISTINGS 1-12 294-41 DIV/FWC
: CURRENT APPLICATION NUMBER: US/08/955,424
: CURRENT FILING DATE: 1997-10-22
: EARLIER APPLICATION NUMBER: 08/568,152
: EARLIER FILING DATE: 1995-06-12
: EARLIER APPLICATION NUMBER: 07/849,420
: EARLIER FILING DATE: 1992-06-24
: EARLIER APPLICATION NUMBER: PCT/NL90/00151
: EARLIER FILING DATE: 1990-10-21
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 5
: LENGTH: 355
: TYPE: PRT
: ORGANISM: Saccharomyces cerevisiae
US-08-955-424-5

Query Match 43.3% Score 38.5; DB 4; Length 355;
Best Local Similarity 50.0%; Pred. No. 3.9e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

Qy 3 SNGNG-----CCNYDG 13
:|||||
Db 184 ASGNGGTRGDCNYDG 199

RESULT 12
US-08-328-961-2
: Sequence 2, Application US/08328961
: Patent No. 5501975
: GENERAL INFORMATION:
: APPLICANT: Chaudhuri, Bhabatosh
: APPLICANT: Stephan, Christine
: APPLICANT: Seeboth, Peter
: APPLICANT: Reizman, Howard
: TITLE OF INVENTION: No. 5501975el DNA Molecules and Hosts
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ciba-Geigy Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: NY
: COUNTRY: USA
: ZIP: 10532
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/462,397
: FILING DATE: 05-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/989,260
: FILING DATE: 11-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Spruill, W. Murray
: REGISTRATION NUMBER: 32,943
: REFERENCE/DOCKET NUMBER: 4-18885/A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8614
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 621 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: APPLICATION NUMBER: US/08/328,961
: FILING DATE:

US-08-462-397-2
```

```
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/989,260
: FILING DATE: 11-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Spruill, W. Murray
: REGISTRATION NUMBER: 32,943
: REFERENCE/DOCKET NUMBER: 4-18885/A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8614
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 621 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-328-961-2

Query Match 43.3% Score 38.5; DB 1; Length 621;
Best Local Similarity 50.0%; Pred. No. 6.5e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

Qy 3 SNGNG-----CCNYDG 13
:|||||
Db 311 ASGNGGTRGDCNYDG 326

RESULT 13
US-08-462-397-2
: Sequence 2, Application US/08462397
: Patent No. 5618690
: GENERAL INFORMATION:
: APPLICANT: Chaudhuri, Bhabatosh
: APPLICANT: Stephan, Christine
: APPLICANT: Seeboth, Peter
: APPLICANT: Reizman, Howard
: TITLE OF INVENTION: No. 5618690el DNA Molecules and Hosts
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ciba-Geigy Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: NY
: COUNTRY: USA
: ZIP: 10532
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/462,397
: FILING DATE: 05-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/989,260
: FILING DATE: 11-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Spruill, W. Murray
: REGISTRATION NUMBER: 32,943
: REFERENCE/DOCKET NUMBER: 4-18885/A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8614
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 621 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: APPLICATION NUMBER: US/08-462-397-2
```

Query Match 43.3%; Score 38.5; DB 1; Length 621;
Best Local Similarity 50.0%; Pred. No. 6.5e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

QY 3 SNGNG-----CCNYDG 13
Db 311 ASGNGGTRGDCNCDYDG 326

RESULT 14
US-08-545-562A-5
; Sequence 5, Application US/08545562A
; Patent No. 5840529
; GENERAL INFORMATION:
; APPLICANT: SEIDAH, Nabil G.
; APPLICANT: DAY, Robert
; APPLICANT: CHRETIEN, Michel
; TITLE OF INVENTION: MAMMALIAN PRO-HORMONE CONVERTASE
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,562A
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/510,347
; FILING DATE: 02-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/517,015
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 200702.90025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 271-3552
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-545-562A-5

Query Match 43.3%; Score 38.5; DB 2; Length 783;
Best Local Similarity 50.0%; Pred. No. 8.1e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

QY 3 SNGNG-----CCNYDG 13
Db 325 ASGNGGQHNDNCNCDYDG 340

RESULT 15
US-09-124-671-23
; Sequence 23, Application US/09124671A
; Patent No. 6160088

; GENERAL INFORMATION:
; APPLICANT: Rothman, James
; APPLICANT: Mayhew, Mark
; APPLICANT: Hoe, Mee
; TITLE OF INVENTION: KDEL RECEPTOR INHIBITORS
; FILE REFERENCE: 31488
; CURRENT APPLICATION NUMBER: US/09/124,671A
; CURRENT FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric human COMP-KDEL
US-09-124-671-23

Query Match 42.7%; Score 38; DB 4; Length 109;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYDGP 14
Db 21 GSSLGGDCCSDLGP 34

Search completed: November 21, 2001, 16:21:21
Job time: 310 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:21:55 ; Search time 26.99 seconds
(without alignments)
39.513 Million cell updates/sec

Title: US-09-443-986A-3

Perfect score: 89

Sequence: 1 GTSNGNGCCNYDGP 14

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	53.9	490	2 D71401	probable selenium-
2	45	50.6	470	2 T45816	hypothetical prote
3	45	50.6	595	1 A46373	probable serine/th
4	45	50.6	733	2 H86345	hypothetical prote
5	44	49.4	492	2 S64598	phosphogluconate d
6	43	48.3	258	2 T13263	hypothetical prote
7	42	47.2	198	2 G96838	hypothetical prote
8	42	47.2	357	2 G69393	conserved hypothet
9	42	47.2	3507	2 T34513	hypothetical prote
10	41	46.1	59	2 S25629	myosin alkali high
11	41	46.1	150	1 MOHUA2	myosin alkali high
12	41	46.1	150	1 MORTA2	myosin alkali high
13	41	46.1	150	1 MORBU2	myosin alkali high
14	41	46.1	150	1 MOCHA2	myosin alkali high
15	41	46.1	185	1 MOCHA2	myosin alkali high
16	41	46.1	189	2 T57590	myosin light chain
17	41	46.1	190	1 MORTA1	myosin alkali high
18	41	46.1	192	1 MORBUA	myosin alkali high
19	41	46.1	192	1 MOCHLA	myosin alkali high
20	41	46.1	193	1 MORTA2	myosin alkali high
21	41	46.1	193	1 MOHUA2	myosin alkali high
22	41	46.1	194	1 MOHUA1	myosin alkali high
23	41	46.1	194	1 MOCHLC	myosin alkali high
24	41	46.1	195	1 MOHUV3	myosin alkali high
25	41	46.1	197	1 MOHUA2	myosin alkali high
26	41	46.1	200	1 MORT3V	myosin alkali high
27	41	46.1	218	2 E83575	hypothetical prote
28	41	46.1	417	1 S47793	valine--pyruvate t
29	41	46.1	417	2 E86032	hypothetical prote

ALIGNMENTS

RESULT 1

D71401

probable selenium-binding protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: columbia

C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000

C:Accession: D71401

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D

P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; G

avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B

Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen

erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;

C.; Chalwatzis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t

A:Reference number: A71400; MUID:98121113

A:Accession: D71401

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-490 <BEV>

A:Cross-references: GB:Z97335; NID:g2244747; PID:g2244759

C:Genetics:

A:Map position: 4COP9-4G3845

C:Superfamily: Caenorhabditis elegans hypothetical protein Y37A1B.5

Query Match 53.9%; Score 48; DB 2; Length 490;
Best Local Similarity 64.3%; Pred. No. 13;
Matches 9; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

Oy 3 SNG--NGCCNYDGP 14

||| ||| ||| ||

Db 14 SNGSKGCKYGGP 27

RESULT 2

T45816

hypothetical protein F2809.230 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 28-Jul-2000

C:Accession: T45816

R:Benes, V.; Rechmann, S.; Borkova, D.; Ansoerge, W.; Mewes, H.W.; Lemcke, K.; Mayer,

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23014

A:Accession: T45816

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-470 <BEN>

A:Cross-references: EMBL:AL137080

A:Experimental source: cultivar Columbia; BAC clone F2809

C:Genetics:

A:Map position: 3

A: Introns: 33/1; 59/1
A: Note: F2809.230
C: Superfamily: Arabidopsis thaliana hypothetical protein F2809.230

Query Match 50.6%; Score 45; DB 2; Length 470;
Best Local Similarity 61.5%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTSNGGCCNYDG 13
| | | | | | | | | |
Db 55 GTSNGTCCDRFG 67

RESULT 3
A46373
probable serine/threonine-specific protein kinase (EC 2.7.1.1-) PRO25 - Arabidopsis thaliana
C: Species: Arabidopsis thaliana (mouse-ear cress)
C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C: Accession: A46373
R: Kohorn, B.D.; Lane, S.; Smith, T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 10989-10992, 1992
A: Title: An Arabidopsis serine/threonine kinase homologue with an epidermal growth factor
A: Reference number: A46373; MUID: 93066369
A: Accession: A46373
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-595 <KOH>
A: Cross-references: GB:L04999; NID: g166812; PIDN: AAA32844.1; PID: g166813
A: Note: Sequence extracted from NCBI backbone (NCBIN: 118788, NCBIPI: 118790)
C: Superfamily: Arabidopsis probable serine/threonine-specific protein kinase PRO25; protein
C: Keywords: phosphotransferase
F: 268-546/Domain: protein kinase homology <KIN>

Query Match 50.6%; Score 45; DB 1; Length 595;
Best Local Similarity 58.3%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GTSNGGCCNYD 12
| | | | | | | | | |
Db 20 GECNGGCCCRVD 31

RESULT 4
H86345
hypothetical protein AAF81356.1 [imported] - Arabidopsis thaliana
C: Species: Arabidopsis thaliana (mouse-ear cress)
C: Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C: Accession: H86345
R: Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A: Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A: Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A: Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A: Reference number: A86141; MUID: 21016719
A: Accession: H86345
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-733 <STO>
A: Cross-references: GB:AE005172; NID: g8920634; PIDN: AAF81356.1; GSPDB: GN00141
C: Genetics:
A: Map position: 1
C: Superfamily: Arabidopsis probable serine/threonine-specific protein kinase PRO25; protein
Query Match 50.6%; Score 45; DB 2; Length 733;
Best Local Similarity 58.3%; Pred. No. 47;

Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 GTSNGGCCNYD 12
| | | | | | | | | |
Db 157 GECNGGCCCRVD 168

RESULT 5
S64588
phosphogluconate dehydrogenase (decarboxylating) (EC 1.1.1.44) 2 - yeast (Saccharomyc
N: Alternate names: protein G9170; protein YGR256w
C: Species: Saccharomyces cerevisiae
C: Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 20-Jun-2000
C: Accession: S64588
R: Agostoni Carbone, M.L.; Panzeri, L.; Melchiorretto, P.; Carignani, G.; Feroli, F.; F
submitted to the Protein Sequence Database, May 1996
A: Reference number: S64577
A: Accession: S64588
A: Molecule type: DNA
A: Residues: 1-492 <AGO>
A: Cross-references: EMBL: Z73041; NID: g1323465; PIDN: CAA97285.1; PID: g1323466; MIPS: YG
A: Experimental source: strain S288C
C: Genetics:
A: Gene: SGD: GND2
A: Cross-references: SGD: S0003488; MIPS: YGR256w
A: Map position: 7R
C: Superfamily: phosphogluconate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate
C: Keywords: oxidoreductase; pentose phosphate pathway
F: 8-287/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>

Query Match 49.4%; Score 44; DB 2; Length 492;
Best Local Similarity 58.3%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 SNGNGCCNYDGP 14
| | | | | | | | | |
Db 166 SNGEPCCEWVG 177

RESULT 6
T13263
hypothetical protein - Lactococcus lactis phage BK5-T
C: Species: Lactococcus lactis phage BK5-T
C: Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999
C: Accession: T13263
R: Boyce, J.D.; Davidson, B.E.; Hillier, A.J.
Appl. Environ. Microbiol. 61, 4089-4098, 1995
A: Title: Sequence analysis of the temperate Lactococcus lactis bacteriophage BK5-T an
A: Reference number: Z17646; MUID: 96064422
A: Accession: T13263
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-258 <BOY>
A: Cross-references: EMBL: L44593; NID: g928826; PID: g928835; PIDN: AAA98586.1

Query Match 48.3%; Score 43; DB 2; Length 258;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GTSNGGCCNYD 12
| | | | | | | | | |
Db 183 GDSNGNGGTNYN 194

RESULT 7
G96838
hypothetical protein T21F11.2 [imported] - Arabidopsis thaliana
C: Species: Arabidopsis thaliana (mouse-ear cress)
C: Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C: Accession: G96838
R: Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: G96838
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-198 <STO>
 A:Cross-references: GB:AE005173; NID:g6730736; PIDN:AAF27126.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T21F11.2
 A:Map position: 1

Query Match 47.2%; Score 42; DB 2; Length 198;
 Best Local Similarity 85.7%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 8 CCNYDGP 14
 ||| |||
 Db 135 CCNYDGP 141

RESULT 8
 G69393
 conserved hypothetical protein AF1152 - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 08-Oct-1999
 C:Accession: G69393
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343
 A:Accession: G69393
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-357 <KLE>
 A:Cross-references: GB:AE001024; GB:AE000782; NID:Q2689347; PIDN:AAB90092.1; PID:g264943
 C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0041

Query Match 47.2%; Score 42; DB 2; Length 357;
 Best Local Similarity 53.3%; Pred. No. 69;
 Matches 8; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
 QY 1 GTSNGNGC--CNYDGS 13
 | ||| ||| |||
 Db 140 GYCNGEGCEVCNFTG 154

RESULT 9
 T34513
 hypothetical protein ZK783.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34513
 R;Favella, A.; Vaudin, M.
 submitted to the EMBL Data Library, August 1994
 A:Description: The sequence of C. elegans cosmid ZK783.
 A:Reference number: Z21536
 A:Accession: T34513
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-3507 <AV>
 A:Cross-references: EMBL:U13646; PIDN:AAC24418.1; GSPDB:GN00021; CESP:ZK783.1
 A:Experimental source: strain Bristol N2; clone ZK783
 C:Genetics:
 A:Gene: CESP:ZK783.1
 A:Map position: 3
 A:Introns: 14/1; 48/2; 84/1; 196/3; 303/1; 381/1; 586/1; 605/1; 1175/3; 1207/1; 1409/
 3504/1

Query Match 47.2%; Score 42; DB 2; Length 3507;
 Best Local Similarity 58.3%; Pred. No. 4.8e+02;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 TSNNGGCCNYDG 13
 | : | | | | | | |
 Db 90 TCSGKGACLYDG 101

RESULT 10
 S25629
 myosin alkali light chain, ventricular - mouse (fragment)
 N:Alternate names: myosin catalytic light chain, ventricular
 C:Species: Mus musculus (house mouse)
 C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 21-Jan-2000
 C:Accession: S25629
 R;Barton, P.; Cohen, A.; Alonso, S.
 submitted to the EMBL Data Library, August 1992
 A:Reference number: S25629
 A:Accession: S25629
 A:Molecule type: mRNA
 A:Residues: 1-59 <BAR>
 A:Cross-references: EMBL:X67685; NID:953299; PIDN:CAA47917.1; PID:g53300
 C:Superfamily: calmodulin; calmodulin repeat homology
 C:Keywords: calcium binding; cardiac muscle; duplication; EF hand; heart; muscle cont
 F:27-59/Domain: calmodulin repeat homology <EFH>

Query Match 46.1%; Score 41; DB 2; Length 59;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GTSNGGCCNYD 12
 | : | | | | |
 Db 39 GQEDSNGGCINYE 50

RESULT 11
 MOHUA2
 myosin alkali light chain 1, fast skeletal muscle, form 2 - human
 N:Alternate names: MLC3fast; myosin A2 catalytic light chain; myosin L3 catalytic lig
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1991 #sequence_revision 03-Oct-1995 #text_change 22-Jun-1999
 C:Accession: S07939
 R;Seidel, U.; Bober, E.; Winter, B.; Lenz, S.; Lohse, P.; Arnold, H.H.
 Nucleic Acids Res. 15, 4989, 1987
 A:Title: The complete nucleotide sequences of cDNA clones coding for human myosin lig
 A:Reference number: S07939; MUID:87259977
 A:Accession: S07939
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-150 <SEI>
 A:Cross-references: EMBL:X05451; NID:g34688; PIDN:CAA29020.1; PID:g34689
 A:Experimental source: fetal skeletal muscle
 C:Genetics:
 A:Gene: GDB:MYL1
 A:Cross-references: GDB:120217; OMIM:160780
 A:Map position: 2q32.1-2qter
 C:Superfamily: calmodulin; calmodulin repeat homology
 C:Keywords: acetylated amino end; alternative splicing; calcium binding; duplication;
 F:6-38/Domain: calmodulin repeat homology <EF1>
 F:83-115/Domain: calmodulin repeat homology <EF3>
 F:118-150/Domain: calmodulin repeat homology <EF4>

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted

Query Match 46.1%; Score 41; DB 1; Length 150;

Best Local Similarity 50.0%; Pred. No. 46;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12

I : I I I I I :

DB 130 GOEDSNGCINYE 141

RESULT 12

MORTA2

myosin alkali light chain 1, fast skeletal muscle, form 2 [similarity] - rat

N:Alternate names: myosin A2 catalytic light chain; myosin light chain 3

C:Species: Rattus norvegicus (Norway rat)

C:Date: 28-Feb-1986 #sequence_revision 16-Jun-2000 #text_change 16-Jun-2000

C:Accession: I77418; A03035

R:Strehler, E.E.; Perlasamy, M.; Strehler-Page, M.

Mol. Cell. Biol. 5, 3168-3182, 1985

A:Title: Myosin light-chain 1 and 3 gene has two structurally distinct and differentially

A:Reference number: I57590; MUID:86310786

A:Accession: I77418

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-150 <RES>

A:Cross-references: GB:M12021; NID:g205482; PIDN:AAA1623.1; PID:g205486

R:Perlasamy, M.; Strehler, E.E.; Garfinkel, L.I.; Gubits, R.M.; Ruiz-Opazo, N.; Nadal-Gi

J. Biol. Chem. 259, 13595-13604, 1984

A:Title: Fast skeletal muscle myosin light chains 1 and 3 are produced from a single gen

A:Reference number: A32456; MUID:85030494

A:Accession: A03035

A:Molecule type: mRNA

A:Residues: 2-113, T', 115-150 <PER>

A:Cross-references: GB:L00088; GB:J00753; NID:g205472; PIDN:AAA98534.1; PID:g205475

A>Note: Initiator Met not shown

C:Genetics:

A:Introns: 1/3; 10/1; 58/1; 116/1; 142/1

C:Superfamily: calmodulin; calmodulin repeat homology

F:2-150/Product: myosin light chain #status predicted <MAT>

F:6-38/Domain: calmodulin repeat homology <EF1>

F:83-115/Domain: calmodulin repeat homology <EF2>

F:118-150/Domain: calmodulin repeat homology <EF3>

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted

Query Match

Best Local Similarity 46.1%; Score 41; DB 1; Length 150;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12

I : I I I I I :

DB 130 GOEDSNGCINYE 141

RESULT 13

MORBU2

myosin alkali light 1, fast skeletal muscle, form 2 - rabbit

N:Alternate names: MLC3fast; myosin A2 catalytic light chain; myosin L3 catalytic light

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 30-Sep-1993 #sequence_revision 03-Oct-1995 #text_change 22-Jun-1999

C:Accession: S12693; B03036

R:Mueller, B.; Maeda, K.; Wittinghofer, A.

Nucleic Acids Res. 18, 6688, 1990

A:Title: Sequence of the myosin light chain 1/3 isolated from a rabbit fast skeletal mus

A:Reference number: S12692; MUID:91067463

A:Accession: S12693

A:Molecule type: mRNA

A:Residues: 1-150 <MUE>

A:Cross-references: EMBL:X54044; NID:g1638; PIDN:CAA37977.1; PID:g1639

R:Frank, G.; Weeds, A.G.

Eur. J. Biochem. 44, 317-334, 1974

A:Title: The amino-acid sequence of the alkali light chains of rabbit skeletal-muscle

A:Reference number: A91219; MUID:74257123

A:Accession: B03036

A:Molecule type: protein

A:Residues: 2-55, D', 57, Q', 59-99, DTVG', 104-150 <FRA>

C:Superfamily: calmodulin; calmodulin repeat homology

C:Keywords: acetylated amino end; alternative splicing; calcium binding; duplication;

F:6-38/Domain: calmodulin repeat homology <EF1>

F:83-115/Domain: calmodulin repeat homology <EF3>

F:118-150/Domain: calmodulin repeat homology <EF4>

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental

Query Match

Best Local Similarity 46.1%; Score 41; DB 1; Length 150;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12

I : I I I I I :

DB 130 GOEDSNGCINYE 141

RESULT 14

MOCHA2

myosin alkali light chain 1, fast skeletal muscle, splice form 2 - chicken

N:Alternate names: myosin A2 catalytic light chain; myosin L4 catalytic light chain

C:Species: Gallus gallus (chicken)

C:Date: 01-Sep-1981 #sequence_revision 16-Oct-1998 #text_change 16-Jun-2000

C:Accession: I50384; I50665; B93443; A91107

R:Nabeshima, Y.; Fujii-Kuriyama, Y.; Muramatsu, M.; Ogata, K.

Nature 308, 333-338, 1984

A:Title: Alternative transcription and two modes of splicing result in two myosin lig

A:Reference number: I50383; MUID:84168118

A:Accession: I50384

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-74, M', 75-103, 105-150 <NAB1>

A:Cross-references: GB:K02610; NID:g212328; PIDN:AAA48956.1; PID:g212331

A:Accession: I50665

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-9 <NAB2>

A:Cross-references: GB:X00459; GB:K02609; NID:g63623; PID:g584508; PID:g1334736

R:Nabeshima, Y.; Fujii-Kuriyama, Y.; Muramatsu, M.; Ogata, K.

Nucleic Acids Res. 10, 6099-6110, 1982

A:Title: Molecular cloning and nucleotide sequences of the complementary DNAs to chic

A:Reference number: A93443; MUID:83064541

A:Accession: B93443

A:Molecule type: mRNA

A:Residues: 2-150 <NAB3>

A:Cross-references: GB:J00888; NID:g212348

R:Malta, T.; Umegane, T.; Matsuda, G.

Eur. J. Biochem. 114, 45-49, 1981

A:Title: Amino-acid sequence of the L-4 light chain of chicken skeletal-muscle myosin

A:Reference number: A91107; MUID:81164496

A:Accession: A91107

A:Molecule type: protein

A:Residues: 2-6, Q', 8, D', 10-99, D', 101-150 <MA1>

C:Genetics:

A:Gene: MLC

A:Introns: 1/3; 10/1; 58/1; 116/1; 142/1

C:Superfamily: calmodulin; calmodulin repeat homology

C:Keywords: alternative splicing; blocked amino end; calcium binding; duplication; EF

F:2-150/Product: myosin alkali light chain 1, fast skeletal muscle, splice form 2 #st

F:6-38/Domain: calmodulin repeat homology <EF1>

F:83-115/Domain: calmodulin repeat homology <EF3>

F:118-150/Domain: calmodulin repeat homology <EF4>

F:2/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #st

Query Match

Best Local Similarity 46.1%; Score 41; DB 1; Length 150;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGGCCNYD 12
| : ||| ||:
Db 130 QEDSNGCINYE 141

RESULT 15

MOCH4E

myosin alkali light chain 4, embryonic - chicken
N;Alternate names: myosin L23 catalytic light chain
C;Species: Gallus gallus (chicken)
C;Date: 01-Dec-1989 #sequence_revision 03-Oct-1995 #text_change 22-Jun-1999
C;Accession: S02065; A29473
R;Nabeshima, Y. I.; Nabeshima, Y.; Kawashima, M.; Nakamura, S.; Nonomura, Y.; Fujii-Kuriyama, Y. Mol. Biol. 204, 497-505, 1988
A;Title: Isolation of the chick myosin alkali light chain gene expressed in embryonic g
A;Reference number: S02065; MUID:89141751
A;Accession: S02065
A;Molecule type: DNA
A;Residues: 1-185 <NAB>
A;Cross-references: EMBL:X14428
A;Note: the authors translated the codon CGG for residue 71 as Ala
R;Kawashima, M.; Nabeshima, Y.; Obinata, T.; Fujii-Kuriyama, Y. J. Biol. Chem. 262, 14408-14414, 1987
A;Title: A common myosin light chain is expressed in chicken embryonic skeletal, cardiac
A;Reference number: A29473; MUID:88032983
A;Accession: A29473
A;Molecule type: mRNA
A;Residues: 1-93, 'L', 95-165, 'L', 167-185 <KAW>
A;Cross-references: GB:J02823; NID:g212339; PIDN:AAA48957.1; PID:g212340
A;Note: the authors translated the codon CTG for residue 94 as Met and TAC for residue 1
C;Comment: In chicken, this form is transiently expressed in embryonic skeletal, cardiac
C;Genetics:
A;Introns: 33/3; 43/1; 93/1; 151/1; 177/1
C;Complex: The myosin molecule contains two heavy chains, two alkali light chains, and t
C;Superfamily: calmodulin; calmodulin repeat homology
C;Keywords: brain; calcium binding; duplication: EF hand; muscle contraction
F;35-73/Domain: calmodulin repeat homology <EF1>
F;118-150/Domain: calmodulin repeat homology <EF3>
F;153-185/Domain: calmodulin repeat homology <EF4>

Query Match 46.1%; Score 41; DB 1; Length 185;
Best Local Similarity 50.0%; Pred. No. 55;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGGCCNYD 12
| : ||| ||:
Db 165 QEDANGCINYE 176

Search completed: November 21, 2001, 16:21:55
Job time: 309 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:28:03 ; Search time 17.74 Seconds
(without alignments)
27.034 Million cell updates/sec

Title: US-09-443-986A-3
Perfect score: 89
Sequence: 1 GTSNGNGCCNYDGP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	48	53.9	490	1 SBP_ARATH	O23284 arabidopsis
2	44	49.4	492	1 6PG2_YEAST	P53319 saccharomyc
3	41	46.1	149	1 MLE3_CHICK	P02605 gallus gall
4	41	46.1	149	1 MLE3_HUMAN	P06741 homo sapien
5	41	46.1	149	1 MLE3_MOUSE	P05978 mus musculu
6	41	46.1	149	1 MLE3_RABIT	P02603 oryctolagus
7	41	46.1	149	1 MLE3_RAT	P02601 rattus norv
8	41	46.1	185	1 MLE3_CHICK	P09540 gallus gall
9	41	46.1	187	1 MLE1_MOUSE	P05977 mus musculu
10	41	46.1	188	1 MLE1_RAT	P02600 rattus norv
11	41	46.1	190	1 MLE1_CHICK	P02604 gallus gall
12	41	46.1	191	1 MLE1_RABIT	P02602 oryctolagus
13	41	46.1	192	1 MLEF_MOUSE	P09541 mus musculu
14	41	46.1	192	1 MLEF_RAT	P17209 rattus norv
15	41	46.1	193	1 MLE1_HUMAN	P05976 homo sapien
16	41	46.1	193	1 MLEC_CHICK	P02606 gallus gall
17	41	46.1	194	1 MLEV_HUMAN	P08590 homo sapien
18	41	46.1	196	1 MLEF_HUMAN	P12829 homo sapien
19	41	46.1	199	1 MLEV_RAT	P16409 rattus norv
20	41	46.1	202	1 T4S4_HUMAN	P48230 homo sapien
21	41	46.1	417	1 AVTA_ECOLI	P09053 escherichia
22	41	46.1	617	1 RNCL_YEAST	P33753 saccharomyc
23	41	46.1	667	1 DIF_DROME	P98149 drosophila
24	41	46.1	1358	1 XDH_CHICK	P47990 gallus gall
25	40	44.9	79	1 KRHA_RABIT	Q02957 oryctolagus
26	40	44.9	141	1 MLEN_MOUSE	Q60605 mus musculu
27	40	44.9	148	1 VEGH_ORFN7	P52585 orf virus (
28	40	44.9	150	1 MLEN_HUMAN	P08296 gallus gall
29	40	44.9	150	1 MLEN_HUMAN	P16475 homo sapien
30	40	44.9	150	1 MLES_CHICK	P02607 gallus gall
31	40	44.9	150	1 MLES_HUMAN	P24572 homo sapien
32	40	44.9	150	1 MLES_RAT	Q64119 rattus norv
33	40	44.9	208	1 MLEV_HUMAN	P14649 homo sapien

34	40	44.9	721	1 ZW10_DROME	Q9w4x9 drosophila
35	40	44.9	2193	1 POLG_HE71B	Q66478 human enter
36	40	44.9	2193	1 POLG_HE71M	Q66479 human enter
37	39.5	44.4	156	1 YQJY_BACSU	P54562 bacillus su
38	39	43.8	88	1 PER_DROME	Q26287 drosophila
39	39	43.8	118	1 LTP2_ARATH	Q9s713 arabidopsis
40	39	43.8	173	1 EGG1_SCHMA	P06649 schistosoma
41	39	43.8	177	1 EGG2_SCHMA	P12796 schistosoma
42	39	43.8	177	1 EGG3_SCHMA	P13396 schistosoma
43	39	43.8	325	1 YA22_SCHPO	Q09695 schizosacch
44	39	43.8	431	1 PRIS_THEMA	Q9x004 thermotoga
45	39	43.8	462	1 PGLR_PERA	Q02096 persea amer

ALIGNMENTS

RESULT 1
SBP_ARATH STANDARD; PRT; 490 AA.
AC O23264;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE PUTATIVE SELENIUM-BINDING PROTEIN.
GN A74G14030 OR DL3055C.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=98121113; PubMed=9461215;
RA Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,
RA Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,
RA Ridley P., Hudson S.-A., Patel K., Murphy G., Piffanelli P.,
RA Wedler H., Wedler E., Wambutt R., Weitzenecker T., Pohl T.M.,
RA Terryn N., Gielen J., Villarroel R., de Clerck R., van Montagu M.,
RA Lecharny A., Aubourg S., Gy I., Kreis M., Lao N., Kavanagh T.,
RA Hempel S., Kottler S., Entian K.-D., Rieger M., Schaeffer M., Funk B.,
RA Mueller-Auer S., Silvey M., James R., Montfort A., Pons A.,
RA Puigdomenech P., Douka A., Voukelatou E., Milioni D., Hatzopoulos P.,
RA Piravandi E., Obermaier B., Hilbert H., Duesterhoeft A., Moores T.,
RA Jones J.D.G., Eneva T., Palme K., Benes V., Rechman S., Ansoorge W.,
RA Cooke R., Berger C., Delsen M., Voet M., Volckaert G., Mewes H.-W.,
RA Klosterman S., Schueller C., Chaltatzis N.
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
Arabidopsis thaliana".
RL Nature 391:485-488(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansoorge W., Brandt P., Givell L., Rieger M., Mueller M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Muller M.,
RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J., Vandenbusche F.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Bastiaens I., Aert R., Defoor E.,
RA Braeken M., Welljens I., Voet M., Ramsperger U., Hilbert H., Braun M.,
RA Weitzenecker T., Bothe G., Ransperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., de Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McMay R.,
RA Pettett A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,

RA Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Farmann B., Granderath K., Dauner A., Herzl A.,
RA Neumann S., Argirioy F., Vitale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lechary A., Aubourg S.,
RA Chedok F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenberg M., Barges M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bieleke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Farnell L., Dedha N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Stokem M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stenking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C.,
RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.,
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
thaliana".
RL Nature 402:769-777(1999).
CC -!- SIMILARITY: BELONGS TO THE SELENIUM-BINDING PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z97335; CAB10182.1; -;
DR EMBL; AL161537; CAB78445.1; -;
KW Selenium.
SQ SEQUENCE 490 AA; 54057 MW; 10EE7B9BCF2F0390 CRC64;

Query Match 53.9%; Score 48; DB 1; Length 490;
Best Local Similarity 64.3%; Pred. No. 4.9;
Matches 9; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

Qy 3 SNG--NGCCNDGDP 14
Db 14 SNGSGKGCCKYGGP 27

RESULT 2
6PG2_YEAST STANDARD; PRT; 492 AA.
AC P53319;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 6-PHOSPHOGLUCONATE DEHYDROGENASE, DECARBOXYLATING 2 (EC 1.1.1.44).
GN GN2 OR YGR256W OR G9170.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=9729233; PubMed=9133741;
RA Mazzoni C., Ruzzi M., Rinaldi T., Solinas F., Montebove F.,
RA Frontali L.;
RT "Sequence analysis of a 10.5 kb DNA fragment from the yeast

RT chromosome VII reveals the presence of three new open reading frames
and of a tRNA^{Thr} gene".
RL Yeast 13:369-372(1997).
CC -!- CATALYTIC ACTIVITY: 6-PHOSPHO-D-GLUCONATE + NADP(+) = D-RIBULOSE
CC 5-PHOSPHATE + CO(2) + NADPH.
CC -!- PATHWAY: HEXOSE MONOPHOSPHATE SHUNT.
CC -!- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X99228; CAA67612.1; -;
DR EMBL; Z73041; CAA97285.1; -;
DR HSSP; P00349; 1PGQ.
DR SGD; S0003488; GND2.
DR InterPro; IPR001744; -;
DR Pfam; PF00393; 6PGD; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
DR PROSITE; PS00461; 6PGD; 1.
KW Oxidoreductase; Pentose shunt; NADP.
SQ SEQUENCE 492 AA; 53922 MW; 3D75D53563987735 CRC64;

Query Match 49.4%; Score 44; DB 1; Length 492;
Best Local Similarity 58.3%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 SNGNGCCNYDGP 14
Db 166 SNGEPCCWVGVP 177

RESULT 3
MLE3_CHICK
ID MLE3_CHICK STANDARD; PRT; 149 AA.
AC P02605;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE MYOSIN LIGHT CHAIN 3, SKELETAL MUSCLE ISOFORM (A2 CATALYTIC) (ALKALI
DE (MLC-3)).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=81164496; PubMed=6783403;
RA Maita T., Umegane T., Matsuda G.;
RT "Amino-acid sequence of the L-4 light chain of chicken
skeletal-muscle myosin".
RL Eur. J. Biochem. 114:45-49(1981).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83064541; PubMed=6128725;
RA Nabeshima Y., Fujii-Kuriyama Y., Muramatsu M., Ogata K.;
RT "Molecular cloning and nucleotide sequences of the complementary DNAs
to chicken skeletal muscle myosin two alkali light chain mRNAs".
RL Nucleic Acids Res. 10:6099-6110(1982).
CC -!- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
CC -!- ALTERNATIVE PRODUCTS: MYOSIN MLC1 & MLC3 LIGHT CHAINS ARE PRODUCED
CC FROM A SINGLE GENE BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
CC PROTEIN DOES NOT BIND CALCIUM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; J00888; AAA48961.1; -;
DR PIR; A03037; MOCHLA.
DR HSP; P02593; 1CDM.
DR InterPro: IPR002048; -;
DR Pfam; PF00036; ehand; 2.
KW Myosin; Muscle protein; Alternative splicing.
FT INIT_MET 0 0
FT DOMAIN 1 8 SPECIFIC TO MLC3.
FT DOMAIN 9 149 COMMON TO MLC1 AND MLC3.
FT CONFLICT 6 6 Q -> E (IN REF. 2).
FT CONFLICT 8 8 D -> N (IN REF. 2).
FT CONFLICT 127 127 M -> I (IN REF. 2).
SQ SEQUENCE 149 AA; 16579 MW; BD083C3EC9D02B90 CRC64;

Query Match 46.1%; Score 41; DB 1; Length 149;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12
| : |||||
Db 129 GQEDSNGCINYE 140

RESULT 4

MLE3_HUMAN
ID MLE3_HUMAN STANDARD; PRT; 149 AA.
AC P06741;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE MYOSIN LIGHT CHAIN 3, SKELETAL MUSCLE ISOFORM (A2 CATALYTIC) (ALKALI
DE (MLC3F).
GN MYL1.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88329722; PubMed=2458299;
RA Seidel U., Bober E., Winter B., Lenz S., Lohse P., Goedde H.,
RA Grzeschik K., Arnold H.H.;
RT "Alkali myosin light chains in man are encoded by a multigene family
RT that includes the adult skeletal muscle, the embryonic or atrial, and
RT nonsarcomeric isoforms.";
RL Gene 66:135-146(1988).

[2]
RP SEQUENCE FROM N.A.
RC TISSUE=fetal muscle;
RX MEDLINE=8725977; PubMed=3601661;
RA Seidel U., Bober E., Winter B., Lenz S., Lohse P., Arnold H.H.;
RT "The complete nucleotide sequences of cDNA clones coding for human
RT myosin light chains 1 and 3.";
RL Nucleic Acids Res. 15:4989-4989(1987).
CC -1- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
CC -1- ALTERNATIVE PRODUCTS: MYOSIN MLC1 & MLC3 LIGHT CHAINS ARE PRODUCED
CC FROM A SINGLE GENE BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
CC PROTEIN DOES NOT BIND CALCIUM.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M20643; AAA59855.1; -;
DR EMBL; X05451; CAA29020.1; -;
DR PIR; S07939; S07939.
DR HSP; P02593; 1CDM.
DR TIM; 160780; -;
DR InterPro: IPR002048; -;
DR Pfam; PF00036; ehand; 2.
KW Myosin; Muscle protein; Alternative splicing; Multigene family.
FT INIT_MET 0 0
FT DOMAIN 1 8 SPECIFIC TO MLC3.
FT DOMAIN 9 149 COMMON TO MLC1 AND MLC3.
SQ SEQUENCE 149 AA; 16553 MW; CD4BFC7305F92DE9 CRC64;

Query Match 46.1%; Score 41; DB 1; Length 149;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12
| : |||||
Db 129 GQEDSNGCINYE 140

RESULT 5

MLE3_MOUSE
ID MLE3_MOUSE STANDARD; PRT; 149 AA.
AC P05978;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MYOSIN LIGHT CHAIN 3, SKELETAL MUSCLE ISOFORM (A2 CATALYTIC) (ALKALI
DE (MLC3F).
GN MYL1 OR MYLF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85024885; PubMed=6091905;
RA Robert B., Daubas P., Akimenko M.-A., Cohen A., Garner I.,
RA Guenet J.-L., Buckingham M.E.;
RT "A single locus in the mouse encodes both myosin light chains 1 and
RT 3, a second locus corresponds to a related pseudogene.";
RL Cell 39:129-140(1984).
CC -1- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
CC -1- ALTERNATIVE PRODUCTS: MYOSIN MLC1 & MLC3 LIGHT CHAINS ARE PRODUCED
CC FROM A SINGLE GENE BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
CC PROTEIN DOES NOT BIND CALCIUM.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; K02241; AAA39719.1; -;
DR EMBL; K02238; AAA39719.1; JOINED.
DR EMBL; K02239; AAA39719.1; JOINED.
DR EMBL; K02240; AAA39719.1; JOINED.
DR PIR; B23253; B23253.
DR HSP; P02593; 1CDM.
DR SWISS-2DPAGE; P05978; MOUSE.
DR MGD; MGI:97269; Mylf.
KW Myosin; Muscle protein; Alternative splicing; Multigene family.
FT INIT_MET 0 0
FT DOMAIN 1 8 SPECIFIC TO MLC3.

FT DOMAIN 9 149 COMMON TO MLC1 AND MLC3.
SQ SEQUENCE 149 AA; 16468 MW; ACBFD896DB332C47 CRC64;

Query Match 46.1%; Score 41; DB 1; Length 149;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 GTSNNGGCCNYD 12
| : ||| ||:
Db 129 GOEDSNGCINYE 140

RESULT 6
MLE3_RABIT
ID MLE3_RABIT STANDARD; PRT; 149 AA.
AC P02603;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MYOSIN LIGHT CHAIN 3, SKELETAL MUSCLE ISOFORM (A2 CATALYTIC) (ALKALI)
DE (MLC3F).
GN MYL1

OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Fast muscle;

RX MEDLINE=91067463; PubMed=2147476;

RA Mueller B., Maeda K., Wittinghofer A.;

RT "Sequence of the myosin light chain 1/3 isolated from a rabbit fast

skelatal muscle lambda library.";

RL Nucleic Acids Res. 18:6688-6688(1990).

RN [2]

RP SEQUENCE.

RX MEDLINE=74257123; PubMed=4838672;

RA Frank G., Weeds A.G.;

RT "The amino-acid sequence of the alkali light chains of rabbit

skeletal-muscle myosin.";

RL Eur. J. Biochem. 44:317-334(1974).

CC -!- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.

CC -!- ALTERNATIVE PRODUCTS: MYOSIN MLC1 & MLC3 LIGHT CHAINS ARE PRODUCED

FROM A SINGLE GENE BY ALTERNATIVE SPLICING.

CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS

PROTEIN DOES NOT BIND CALCIUM.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X54044; CAA37977.1; .

DR PIR; B03036; MOREL2.

DR PIR; S12693; S12693.

DR HSSP; P02593; ICDM.

DR InterPro; IPR002048; .

DR Pfam; PF00036; ehand; 2.

KW Myosin; Muscle protein; Alternative splicing; Multigene family;

KW Acetylation.

FT INIT_MET 0 0

FT MOD_RES 1 1 ACETYLATION.

FT DOMAIN 1 1 SPECIFIC TO MLC3.

FT DOMAIN 9 149 COMMON TO MLC1 AND MLC3.

FT CONFLICT 55 57 NEE -> DEQ (IN REF. 2).

FT CONFLICT 99 102 NGTV -> DTVG (IN REF. 2).

SQ SEQUENCE 149 AA; 16527 MW; 10F4C3576A8778DC CRC64;

Query Match 46.1%; Score 41; DB 1; Length 149;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 GTSNNGGCCNYD 12
| : ||| ||:
Db 129 GOEDSNGCINYE 140

RESULT 7

MLE3_RAT

ID MLE3_RAT STANDARD; PRT; 149 AA.

AC P02601;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 01-FEB-1996 (Rel. 33, Last annotation update)

DE MYOSIN LIGHT CHAIN 3, SKELETAL MUSCLE ISOFORM (A2 CATALYTIC) (ALKALI)

DE (MLC3F).

GN MYL1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86310786; PubMed=3018505;

RA Strehler E.E., Periasamy M., Strehler-Page M.-A., Nadal-Ginard B.;

RT "Myosin light-chain 1 and 3 gene has two structurally distinct and

differentially regulated promoters evolving at different rates.";

RL Mol. Cell. Biol. 5:3168-3182(1985).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=85030494; PubMed=6092382;

RA Periasamy M., Strehler E.E., Garfinkel L.I., Gubits R.M.,

RT "Fast skeletal muscle myosin light chains 1 and 3 are produced from a

single gene by a combined process of differential RNA transcription

and splicing.";

RL J. Biol. Chem. 259:13595-13604(1984).

CC -!- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.

CC -!- ALTERNATIVE PRODUCTS: MYOSIN MLC1 & MLC3 LIGHT CHAINS ARE PRODUCED

FROM A SINGLE GENE BY ALTERNATIVE SPLICING.

CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS

PROTEIN DOES NOT BIND CALCIUM.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M12021; AAA41623.1; .

DR EMBL; M12018; AAA41623.1; JOINED.

DR EMBL; M12019; AAA41623.1; JOINED.

DR EMBL; M12020; AAA41623.1; JOINED.

DR EMBL; L00088; AAA98534.1; .

DR EMBL; K02424; AAA98534.1; JOINED.

DR EMBL; K02425; AAA98534.1; JOINED.

DR EMBL; L00085; AAA98534.1; JOINED.

DR EMBL; L00086; AAA98534.1; JOINED.

DR EMBL; L00087; AAA98534.1; JOINED.

DR PIR; A03035; MORTA2.

DR HSSP; P02593; ICDM.

KW Myosin; Muscle protein; Alternative splicing; Multigene family.

FT INIT_MET 0 0

FT DOMAIN 1 8 SPECIFIC TO MLC3.

FT DOMAIN 9 149 COMMON TO MLC1 AND MLC3.

FT CONFLICT 113 113 T -> I (IN REF. 1).

SQ SEQUENCE 149 AA; 16483 MW; ACBFD94DABAD5D02 CRC64;

Query Match 46.1%; Score 41; DB 1; Length 149;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTSNGNGCCNYD 12
 | : ||| ||:
 Db 129 GQEDSNGCINYE 140

RESULT 8
 MLEX_CHICK STANDARD; PRT; 185 AA.
 AC P09540; Q90897;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE MYOSIN LIGHT CHAIN, EMBRYONIC (L23).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88032983; PubMed=3667580;
 RA Kawashima M., Nabeshima Y.-I., Obinata T., Fujii-Kuriyama Y.;
 RT "A common myosin light chain is expressed in chicken embryonic
 skeletal, cardiac, and smooth muscles and in brain continuously from
 embryo to adult.";
 RL J. Biol. Chem. 262:14408-14414(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89141751; PubMed=3225843;
 RA Nabeshima Y.I., Nabeshima Y.-I., Kawashima M., Nakamura S.,
 RA Nonomura Y., Fujii-Kuriyama Y.;
 RT "Isolation of the chick myosin alkali light chain gene expressed in
 embryonic gizzard muscle and transitional expression of the light
 chain gene family in vivo.";
 RL J. Mol. Biol. 204:497-505(1988).
 CC -!- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
 CC -!- MISCELLANEOUS: LIGHT CHAIN 23 IS EXPRESSED IN CHICKEN EMBRYONIC
 CC SKELETAL, CARDIAC, SMOOTH MUSCLES AND IN BRAIN CONTINUOUSLY FROM
 CC EMBRYO TO ADULT.
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
 CC PROTEIN DOES NOT BIND CALCIUM.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; J02823; AAA48957.1; -;
 CC EMBL; X14428; CAA32588.1; -;
 CC EMBL; X14429; CAA32588.1; JOINED.
 CC EMBL; M34990; AAA96670.1; ALT_SEQ.
 CC EMBL; M34989; AAA96670.1; JOINED.
 CC PIR; S02065; S02065.
 CC HSSP; P02593; ICDM.
 KW Myosin; Muscle protein; Multigene family.
 FT CONFLICT 94 94 L -> M (IN REF. 2).
 SQ SEQUENCE 185 AA; 20619 MW; 50F87A074F23BACF CRC64;

Query Match 46.1%; Score 41; DB 1; Length 185;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTSNGNGCCNYD 12
 | : ||| ||:
 Db 167 GQEDSNGCINYE 178

Db 165 GLEDANGCINYE 176

RESULT 9
 MLEI_MOUSE STANDARD; PRT; 187 AA.
 AC P05977;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE MYOSIN LIGHT CHAIN 1, SKELETAL MUSCLE ISOFORM (MLC1F) (A1 CATALYTIC)
 DE (ALKALI).
 GN MYL1 OR MYLF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85024885; PubMed=6091905;
 RA Robert B., Daubas P., Akimenko M.-A., Cohen A., Garner I.,
 RA Guenet J.-L., Buckingham M.E.;
 RT "A single locus in the mouse encodes both myosin light chains 1 and
 3, a second locus corresponds to a related pseudogene.";
 RL Cell 39:129-140(1984).
 RN [2]
 RP SEQUENCE OF 1-37 FROM N.A.
 RX MEDLINE=89057447; PubMed=3194193;
 RA Cohen A., Barton P.J.R., Robert B., Garner I., Alonso S.,
 RA Buckingham M.E.;
 RT "Promoter analysis of myosin alkali light chain genes expressed in
 mouse striated muscle.";
 RL Nucleic Acids Res. 16:10037-10052(1988).
 CC -!- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
 CC -!- ALTERNATIVE PRODUCTS: MYOSIN MLC1 & MLC3 LIGHT CHAINS ARE PRODUCED
 CC FROM A SINGLE GENE BY ALTERNATIVE SPLICING.
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
 CC PROTEIN DOES NOT BIND CALCIUM.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; X12973; CAA31416.1; -;
 CC EMBL; K02241; AAA39718.1; -;
 CC EMBL; K02237; AAA39718.1; JOINED.
 CC EMBL; K02238; AAA39718.1; JOINED.
 CC EMBL; K02239; AAA39718.1; JOINED.
 CC EMBL; K02240; AAA39718.1; JOINED.
 CC PIR; A23253; A23253.
 CC PIR; S02259; S02259.
 CC HSSP; P02593; ICDM.
 CC SWISS-2DPAGE; P05977; MOUSE.
 DR MGD; MGI:97269; Mylf.
 KW Myosin; Muscle protein; Alternative splicing; Multigene family.
 FT INIT_MET 0 0
 FT DOMAIN 1 46 SPECIFIC TO MLC1.
 FT DOMAIN 47 187 COMMON TO MLC1 AND MLC3.
 SQ SEQUENCE 187 AA; 20463 MW; 7B35E6019FC9C4CC CRC64;

Query Match 46.1%; Score 41; DB 1; Length 187;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTSNGNGCCNYD 12
 | : ||| ||:
 Db 167 GQEDSNGCINYE 178

```
RESULT 10
MLEL_RAT
ID MLEL_RAT STANDARD; PRT; 188 AA.
AC P02600;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE MYOSIN LIGHT CHAIN 1, SKELETAL MUSCLE ISOFORM (MLC1F) (A1 CATALYTIC)
DE (ALKALI).
GN MYL1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86310786; PubMed=3018505;
RA Strehler E.E., Periasamy M., Strehler-Page M.-A., Nadal-Ginard B.;
RT "Myosin light-chain 1 and 3 gene has two structurally distinct and
RT differentially regulated promoters evolving at different rates.";
RL Mol. Cell. Biol. 5:3168-3182(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85030494; PubMed=6092382;
RA Periasamy M., Strehler E.E., Garfinkel L.I., Gubits R.M.,
RA Ruiz-Opazo N., Nadal-Ginard B.;
RT "Fast skeletal muscle myosin light chains 1 and 3 are produced from a
RT single gene by a combined process of differential RNA transcription
RT and splicing.";
RL J. Biol. Chem. 259:13595-13604(1984).
CC -!- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
CC -!- ALTERNATIVE PRODUCTS: MYOSIN MLC1 & MLC3 LIGHT CHAINS ARE PRODUCED
CC FROM A SINGLE GENE BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
CC PROTEIN DOES NOT BIND CALCIUM.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M12021; AAA41622.1;
DR EMBL; M12017; AAA41622.1; JOINED.
DR EMBL; M12018; AAA41622.1; JOINED.
DR EMBL; M12019; AAA41622.1; JOINED.
DR EMBL; M12020; AAA41622.1; JOINED.
DR EMBL; L00088; AAA98533.1;
DR EMBL; K02423; AAA98533.1; JOINED.
DR EMBL; K02426; AAA98533.1; JOINED.
DR EMBL; L00085; AAA98533.1; JOINED.
DR EMBL; L00086; AAA98533.1; JOINED.
DR EMBL; L00087; AAA98533.1; JOINED.
DR HSP; P02593; ICDM1.
DR MYOSIN; Muscle protein; Alternative splicing; Multigene family.
KW INIT_MET 0
FT DOMAIN 1 47 SPECIFIC TO MLC1.
FT DOMAIN 48 188 COMMON TO MLC1 AND MLC3.
FT CONFLICT 25 25 A -> APA (IN REF. 2).
FT CONFLICT 152 152 T -> I (IN REF. 1).
FT SEQUENCE 188 AA; 20548 MW; 685373ACA28F0B22 CRC64;
Query Match 46.1%; Score 41; DB 1; Length 188;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTSNGNGCCNYD 12
| : ||| |:
```

```
Db 168 GOEDSNGCINYE 179
RESULT 11
MLEL_CHICK
ID MLEL_CHICK STANDARD; PRT; 190 AA.
AC P02604;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE MYOSIN LIGHT CHAIN 1, SKELETAL MUSCLE ISOFORM (A1 CATALYTIC) (ALKALI)
DE (MLC-1).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=81212763; PubMed=7238855;
RA Matsuda G., Maita T., Umegane T.;
RT "The primary structure of L-1 light chain of chicken fast skeletal
RT muscle myosin and its genetic implication.";
RL FEBS Lett. 126:111-113(1981).
RN [2]
RP SEQUENCE.
RX MEDLINE=83081015; PubMed=7173826;
RA Umegane T., Maita T., Matsuda G.;
RT "Amino-acid sequence of the L-1 light chain of chicken fast skeletal-
RT muscle myosin.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1321-1330(1982).
RN [3]
RP SEQUENCE OF 14-190 FROM N.A.
RX MEDLINE=83064541; PubMed=6128725;
RA Nabeshima Y., Fujii-Kuriyama Y., Muramatsu M., Ogata K.;
RT "Molecular cloning and nucleotide sequences of the complementary DNAs
RT to chicken skeletal muscle myosin two alkali light chain mRNAs.";
RL Nucleic Acids Res. 10:6099-6110(1982).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=93303624; PubMed=8316857;
RA Rayment I., Rypniewski W.R., Schmidt-Base K., Smith R.,
RA Tomchick D.R., Benning M.M., Winkelmann D.A., Wesenberg G.,
RA Holden H.M.;
RT "Three-dimensional structure of myosin subfragment-1: a molecular
RT motor ";
RL Science 261:50-58(1993).
CC -!- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
CC -!- ALTERNATIVE PRODUCTS: MYOSIN MLC1 & MLC3 LIGHT CHAINS ARE PRODUCED
CC FROM A SINGLE GENE BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
CC PROTEIN DOES NOT BIND CALCIUM.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; J00887; AAA48960.1;
DR PIR; A03037; MOCHLA.
DR PDB; 2MYS; 11-JAN-97.
DR InterPro; IPR002048;
DR Pfam; PF00036; ehand; 2.
KW MYOSIN; Muscle protein; Alternative splicing; 3D-structure.
FT MOD_RES 1 1 BLOCKED.
FT DOMAIN 1 49 SPECIFIC TO MLC1.
FT DOMAIN 50 190 COMMON TO MLC1 AND MLC3.
FT CONFLICT 44 44 S -> F (IN REF. 3).
FT SEQUENCE 190 AA; 20696 MW; AD793A0DD9E399F7 CRC64;
```

```

Query Match      46.1%; Score 41; DB 1; Length 190;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 1 GTSNGNGCCNYD 12
   | : ||| ||:
Db 170 GQEDSNGCINYE 181

RESULT 12
MLEI_RABIT
ID MLEI_RABIT STANDARD; PRT; 191 AA.
AC P02602;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MYOSIN LIGHT CHAIN 1, SKELETAL MUSCLE ISOFORM (MLC1F) (AI CATALYTIC)
DE (ALKALI).
GN MYL1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fast muscle;
RX MEDLINE=91067463; PubMed=2147476;
RA Mueller B., Maeda K., Wittinghofer A.;
RT "Sequence of the myosin light chain 1/3 isolated from a rabbit fast
skeletal muscle lambda library.";
RL Nucleic Acids Res. 18:6688-6688(1990).
RN [2]
RP SEQUENCE OF 2-191.
RX MEDLINE=74257123; PubMed=4838672;
RA Frank G., Weeds A.G.;
RT "The amino-acid sequence of the alkali light chains of rabbit
skeletal-muscle myosin.";
RL Eur. J. Biochem. 44:317-334(1974).
RN [3]
RP REVISIONS TO 140-141.
RX MEDLINE=76210734; PubMed=1227935;
RA Weeds A.G.;
RT "Cyanogen bromide fragments of the cardiac I light chain from bovine
myosin: evidence for sequence homology with rabbit skeletal myosin
alkali light chains.";
RL FEBS Lett. 59:203-208(1975).
RN [4]
RP IDENTIFICATION OF N-TERMINAL RESIDUE.
RA Henry G.D., Dalgarno D.C., Levine B.A., Trayer I.P.;
RT "Discovery of alpha-N-trimethylalanine in myosin light chains and its
role in actomyosin interaction.";
RL Biochem. Soc. Trans. 10:362-363(1982).
RN [5]
RP IDENTIFICATION OF N-TERMINAL RESIDUE.
RX MEDLINE=82262099; PubMed=7106295;
RA Henry G.D., Dalgarno D.C., Marcus G., Scott M., Levine B.A.,
RA Trayer I.P.;
RT "The occurrence of alpha-N-trimethylalanine as the N-terminal amino
acid of some myosin light chains.";
RL FEBS Lett. 144:11-15(1982).
CC -!- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
CC -!- ALTERNATIVE PRODUCTS: MYOSIN MLC1 & MLC3 LIGHT CHAINS ARE PRODUCED
FROM A SINGLE GENE BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
PROTEIN DOES NOT BIND CALCIUM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC -----
or send an email to license@isb-sib.ch).
CC -----
EMBL; X54041; CAA37974.1; -.
DR PIR: A03036; MORBLA.
DR PIR: S12692; S12692.
DR PIR: S15061; S15061.
DR HSSP: P02593; LCDM.
DR InterPro: IPR002048; -.
KW Pfam: PF00036; eFhand; 2.
KW Myosin; Muscle protein; Alternative splicing; Multigene family;
KW Methylation.
FT INIT_MET 0 0
FT MOD_RES 1 1 METHYLATION (TRI-).
FT DOMAIN 1 50 SPECIFIC TO MLC1.
FT DOMAIN 51 191 COMMON TO MLC1 AND MLC3.
FT CONFLICT 5 5 D -> N (IN REF. 2).
FT CONFLICT 18 18 A -> KA (IN REF. 2).
FT CONFLICT 29 29 MISSING (IN REF. 2).
FT CONFLICT 97 99 NEE -> DEQ (IN REF. 2).
FT CONFLICT 141 144 NGTV -> DTVG (IN REF. 2).
SQ SEQUENCE 191 AA; 20817 MW; 5D3CEC1730D13E0D CRC64;

Query Match      46.1%; Score 41; DB 1; Length 191;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 1 GTSNGNGCCNYD 12
   | : ||| ||:
Db 171 GQEDSNGCINYE 182

RESULT 13
MLEF_MOUSE
ID MLEF_MOUSE STANDARD; PRT; 192 AA.
AC P09541;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MYOSIN LIGHT CHAIN 1, ATRIAL/PETAL ISOFORM (MLC1A) (MLCIEMB).
GN MLC1A OR MYLA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H;
RX MEDLINE=88315068; PubMed=2842339;
RA Barton P.J.R., Robert B., Cohen A., Garner I., Sassoon D., Weydert A.,
RA Buckingham M.E.;
RT "Structure and sequence of the myosin alkali light chain gene
expressed in adult cardiac atria and fetal striated muscle.";
RL J. Biol. Chem. 263:12669-12676(1988).
RN [2]
RP SEQUENCE OF 1-40 FROM N.A.
RC STRAIN=C3H;
RX MEDLINE=89057447; PubMed=3194193;
RA Cohen A., Barton P.J.R., Robert B., Garner I., Alonso S.,
RA Buckingham M.E.;
RT "Promoter analysis of myosin alkali light chain genes expressed in
mouse striated muscle.";
RL Nucleic Acids Res. 16:10037-10052(1988).
CC -!- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ATRIAL MUSCLE AND IN FETAL
SKELETAL AND VENTRICULAR MUSCLE.
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
PROTEIN DOES NOT BIND CALCIUM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M20772; AAA39721.1; -;
 DR EMBL; M31017; AAA39721.1; JOINED.
 DR EMBL; M20769; AAA39721.1; JOINED.
 DR EMBL; M20770; AAA39721.1; JOINED.
 DR EMBL; M20771; AAA39721.1; JOINED.
 DR EMBL; X12971; CAA31414.1; -;
 DR EMBL; M19436; AAA39720.1; -;
 DR PIR; S01944; S01944.
 DR PIR; A31114; A31114.
 DR HSSP; P02593; ICDM.
 DR MGD; MGI:97267; Myla.
 KW Myosin; Muscle protein; Multigene family.
 FT INIT_MET 0
 SQ SEQUENCE 192 AA; 21028 MW; B0439D512FA19C38 CRC64;

Query Match 46.1%; Score 41; DB 1; Length 192;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNNGCCNYD 12
 I : ||| ||:
 Db 172 GOEDANGCINYE 183

RESULT 14
 MLEF-RAT
 ID MLEF-RAT STANDARD; PRT; 192 AA.
 AC P17209;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-AUG-1990 (Rel. 15, Last annotation update)
 DE MYOSIN LIGHT CHAIN 1, ATRIAL ISOFORM.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Heart;
 RX MEDLINE=902121887; PubMed=2326197;
 RA Rovner A.S., McNally E.M., Leinwand L.A.;
 RT "Complete cDNA sequence of rat atrial myosin light chain 1: patterns
 of expression during development and with hypertension.";
 RL Nucleic Acids Res. 18:1581-1586(1990).
 CC -!- SIMILARITY: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS.
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
 CC PROTEIN DOES NOT BIND CALCIUM.
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X51531; CAA35911.1; -;
 DR PIR; S09236; S09236.
 DR HSSP; P02593; ICDM.
 KW Myosin; Muscle protein; Multigene family.
 FT INIT_MET 0
 SQ SEQUENCE 192 AA; 21151 MW; 25E5E2BA8BC97235 CRC64;

Query Match 46.1%; Score 41; DB 1; Length 192;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNNGCCNYD 12
 I : ||| ||:
 Db 172 GOEDANGCINYE 183

RESULT 15
 MLE1_HUMAN
 ID MLE1_HUMAN STANDARD; PRT; 193 AA.
 AC P05976;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE MYOSIN LIGHT CHAIN 1, SKELETAL MUSCLE ISOFORM (MLC1F) (A1 CATALYTIC)
 DE (ALKALI).
 GN MYL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=88329722; PubMed=2458299;
 RX Seidel U., Bober E., Winter B., Lenz S., Lohse P., Goedde H.,
 RA Grzeschik K., Arnold H.H.;
 RT "Alkali myosin light chains in man are encoded by a multigene family
 that includes the adult skeletal muscle, the embryonic or atrial, and
 nonsarcomeric isoforms.";
 RL Gene 66:135-146(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal muscle;
 RX MEDLINE=87259977; PubMed=3601661;
 RA Seidel U., Bober E., Winter B., Lenz S., Lohse P., Arnold H.H.;
 RT "The complete nucleotide sequences of cDNA clones coding for human
 myosin light chains 1 and 3.";
 RL Nucleic Acids Res. 15:4989-4989(1987).
 RN [3]
 RP SEQUENCE OF 1-43 FROM N.A.
 RX MEDLINE=89380210; PubMed=2777779;
 RA Seidel U., Arnold H.H.;
 RT "Identification of the functional promoter regions in the human gene
 encoding the myosin alkali light chains MLC1 and MLC3 of fast
 skeletal muscle.";
 RL J. Biol. Chem. 264:16109-16117(1989).
 CC -!- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
 CC -!- ALTERNATIVE PRODUCTS: MYOSIN MLC1 & MLC3 LIGHT CHAINS ARE PRODUCED
 FROM A SINGLE GENE BY ALTERNATIVE SPLICING.
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
 CC PROTEIN DOES NOT BIND CALCIUM.
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M20642; AAA59854.1; -;
 DR EMBL; X05450; CAB42646.1; -;
 DR EMBL; J05026; AAA66960.1; -;
 DR PIR; A34249; A34249.
 DR PIR; S07393; S07393.
 DR PIR; JS0431; JS0431.
 DR HSSP; P02593; ICDM.
 DR SWISS-2DPAGE; P05976; HUMAN.
 DR MIN; 160780; -;
 DR InterPro; IPR02048; -;
 DR Pfam; PF00036; efhand; 2
 KW Myosin; Muscle protein; Alternative splicing; Multigene family.
 FT INIT_MET 0
 DOMAIN 1 52
 DOMAIN 53 193
 SPECIFIC TO MLC1.
 COMMON TO MLC1 AND MLC3.

FT CONFLICT 44 44 I -> M (IN REF. 2).
SQ SEQUENCE 193 AA; 21014 MW; 851FF397DA8A8ED4 CRC64;

Query Match 46.18; Score 41; DB 1; Length 193;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12
| : ||| ||:
Db 173 QQEDSNGGCINYE 184

Search completed: November 21, 2001, 16:28:03
Job time: 397 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 16:27:39 ; Search time 46.09 Seconds
(without alignments)
40.188 Million cell updates/sec

Title: US-09-443-986A-3
Perfect score: 89
Sequence: 1 GTSNGNGCCNYDGP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- SPTREMBL_16:*
- 1: sp-archaea:*
 - 2: sp-bacteria:*
 - 3: sp-fungi:*
 - 4: sp-human:*
 - 5: sp-invertebrate:*
 - 6: sp-mammal:*
 - 7: sp-mhc:*
 - 8: sp-organelle:*
 - 9: sp-phase:*
 - 10: sp-plant:*
 - 11: sp-rodent:*
 - 12: sp-unclassified:*
 - 13: sp-vertebrate:*
 - 14: sp-virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	53.9	146	10 Q42178	Q42178 arabidopsis
2	48	53.9	268	5 Q9VHG3	Q9VHG3 drosophila
3	45	50.6	215	5 Q9GQH3	Q9GQH3 drosophila
4	45	50.6	217	5 Q9GQH1	Q9GQH1 drosophila
5	45	50.6	470	10 Q9W2K9	Q9W2K9 arabidopsis
6	45	50.6	595	10 Q39191	Q39191 arabidopsis
7	45	50.6	733	10 Q9LMP0	Q9LMP0 arabidopsis
8	45	50.6	735	10 Q81820	Q81820 arabidopsis
9	45	50.6	790	3 Q13304	Q13304 pneumocysti
10	45	50.6	883	10 Q9W9X1	Q9W9X1 arabidopsis
11	44	49.4	109	5 Q9W2B6	Q9W2B6 drosophila
12	44	49.4	165	13 Q9DFE9	Q9DFE9 oncorhynch
13	43	48.3	258	9 Q38326	Q38326 lactococcus
14	43	48.3	1036	10 Q9FN37	Q9FN37 arabidopsis
15	42	47.2	198	10 Q9W8N2	Q9W8N2 arabidopsis
16	42	47.2	319	5 Q9VRS5	Q9VRS5 drosophila
17	42	47.2	329	5 Q9U8W3	Q9U8W3 brachiost
18	42	47.2	357	1 Q29113	Q29113 archaeoglob
19	42	47.2	520	3 P87105	P87105 pneumocysti

20	42	47.2	623	5 Q9GV80	Q9GV80 drosophila
21	42	47.2	732	5 Q9GTV1	Q9GTV1 drosophila
22	42	47.2	734	5 Q9VOK7	Q9VOK7 drosophila
23	42	47.2	913	5 Q9V7G6	Q9V7G6 drosophila
24	42	47.2	3507	5 Q23587	Q23587 caenorhabdi
25	42	47.2	3550	11 Q9JJN2	Q9JJN2 mus musculu
26	41.5	46.6	1006	2 Q9E2K2	Q9E2K2 pseudomonas
27	41	46.1	32	6 Q9TUJ9	Q9TUJ9 sus scrofa
28	41	46.1	59	11 Q61935	Q61935 mus musculu
29	41	46.1	115	11 Q9R139	Q9R139 spermophilu
30	41	46.1	140	13 Q9PWQ5	Q9PWQ5 gallus gall
31	41	46.1	150	6 Q29069	Q29069 sus scrofa
32	41	46.1	150	13 Q91793	Q91793 xenopus lae
33	41	46.1	150	13 Q90896	Q90896 gallus gall
34	41	46.1	154	5 Q9U6H4	Q9U6H4 drosophila
35	41	46.1	168	6 Q29436	Q29436 bos taurus
36	41	46.1	190	13 Q91792	Q91792 xenopus lae
37	41	46.1	192	13 Q90895	Q90895 gallus gall
38	41	46.1	195	4 Q9NRS8	Q9NRS8 homo sapien
39	41	46.1	218	2 Q915X9	Q915X9 pseudomonas
40	41	46.1	241	10 Q9LE91	Q9LE91 oryza sativ
41	41	46.1	569	3 Q9HGQ5	Q9HGQ5 saccharomys
42	41	46.1	633	10 Q9LZU6	Q9LZU6 arabidopsis
43	41	46.1	667	5 Q9VJE2	Q9VJE2 drosophila
44	41	46.1	711	10 Q9LJW7	Q9LJW7 arabidopsis
45	41	46.1	737	5 Q9GUP9	Q9GUP9 leishmania

ALIGNMENTS

RESULT 1					
Q42178	Q42178	PRELIMINARY:	PRT:	146 AA.	
ID	Q42178				
AC	Q42178				
DT	01-NOV-1996 (TrEMBLrel. 01, Created)				
DR	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)				
DE	01-MAY-2000 (TrEMBLrel. 13, Last annotation update)				
DE	SELENIUM BINDING PROTEIN (FRAGMENT)				
OS	Arabisopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;				
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;				
OC	Brassicales; Brassicaceae; Arabidopsis.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RC	SEQUENCE FROM N.A.				
RP	TISSUE-DRY SEEDS OF A THALIANA ECOTYPE COLUMBIA;				
RA	Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.;				
RL	Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; 227257; CAA81769.1; -				
DR	Mendel; 12761; Arath; 1961; 12761.				
DR	InterPro; IPR000345; -				
DR	PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.				
FT	NON_TER 146				
SQ	SEQUENCE 146 AA; 15446 MW; 5EC67EBC517808D5 CRC64;				

Query Match	53.9%;	Score 48;	DB 10;	Length 146;
Best Local Similarity	64.3%;	Pred. No. 1.2;		
Matches	9;	Conservative	0;	Mismatches
QY	3 SNG--NGCCNYDGP 14			
Db	14 SNGSGKCCRYGGP 27			
RESULT 2				
Q9VHG3	Q9VHG3	PRELIMINARY:	PRT:	268 AA.
ID	Q9VHG3			
AC	Q9VHG3			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DR	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DE	01-MAR-2001 (TrEMBLrel. 16, Last annotation update)			

DE CG12952 PROTEIN.
GN SAGE OR CG12952.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram E.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.B., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AF003682; AAF54351.1; -.
DR HSSP; P10085; IMDY.
DR FlyBase; FBgn037672; CG12952.
DR InterPro; IPR001092; -.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00333; HLH; 1.
SQ SEQUENCE 268 AA; 28638 MW; 229FAF140C9E728C CRC64;

Query Match 53.98; Score 48; DB 5; Length 268;
Best Local Similarity 70.08; Pred. No. 2.1;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 NGNGCCNVDG 13
| | | | | | | |
Db 244 NGGCCAWSG 253

RESULT 3
Q9GQH3 PRELIMINARY; PRT; 215 AA.
AC Q9GQH3:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PERIOD (FRAGMENT).
GN PER.
OS Drosophila teissieri (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=I;
RX MEDLINE=20460389; PubMed=11007323;
RA Lachaise D., Harry M., Solignac M., Lemeunier F., Benassi V.,
RA Cariou M.L.;
RT "Evolutionary novelties in islands: Drosophila santomea, a new
RT melanogaster sister species from Sao Tome.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 267:1487-1495(2000).
DR EMBL; AF251251; AAG44583.1; -.
FT NON_TER 1 215
FT NON_TER 1 215
SQ SEQUENCE 215 AA; 20355 MW; D7F52AB2DF90651C CRC64;

Query Match 50.6%; Score 45; DB 5; Length 215;
Best Local Similarity 61.5%; Pred. No. 5.4;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTSNGGCCNVDG 13
| | | | | | | |
Db 41 GTGNGSGSSNYRG 53

RESULT 4
Q9GQH1 PRELIMINARY; PRT; 217 AA.
AC Q9GQH1:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PERIOD (FRAGMENT).
GN PER.

OS Drosophila teissieri (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4;
RX MEDLINE=20460389; PubMed=11007323;
RA Lachaise D., Harry M., Solignac M., Lemeunier F., Benassi V.,
RA Cariou M.L.;
RT "Evolutionary novelties in islands: Drosophila santomea, a new
RT melanogaster sister species from Sao Tome.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 267:1487-1495(2000).
DR EMBL; AF251253; AAG44585.1; -.
FT NON_TER 1 217
FT NON_TER 1 217
SQ SEQUENCE 217 AA; 20484 MW; E4C4FE3BB79FE0C7 CRC64;

Query Match 50.6%; Score 45; DB 5; Length 217;
Best Local Similarity 61.5%; Pred. No. 5.4;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTSNGGCCNVDG 13
| | | | | | | |
Db 41 GTGNGSGSSNYRG 53

RESULT 5
Q9M2K9 PRELIMINARY; PRT; 470 AA.
ID Q9M2K9

AC Q9W2K9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE HYPOTHEtical 53.5 KDA PROTEIN.
 GN F2809.230.
 OS Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Benes V., Rechmann S., Borkova D., Ansoorge W., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X., Quetier F., Salanoubat M.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL137080; CAB6144.1; -;
 KW Hypothetical protein
 SQ SEQUENCE 470 AA; 53527 MW; F9058A7136AA8066 CRC64;

Query Match 50.6%; Score 45; DB 10; Length 470;
 Best Local Similarity 61.5%; Pred. No. 11;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYDG 13
 ||||| ||:
 Db 55 GTSNGTICCDRTG 67

RESULT 6
 ID Q39191 PRELIMINARY; PRT; 595 AA.
 AC Q39191;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE SERINE THREONINE KINASE.
 GN PRO25.
 OS Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith T.A., Kohorn B.D.;
 RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(0).
 CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; L04999; AAA32844.1; -;
 DR Mendel; 13882; Arath;1197:13882.
 DR InterPro; IPR00152; -;
 DR InterPro; IPR000561; -;
 DR InterPro; IPR000719; -;
 DR InterPro; IPR001881; -;
 DR InterPro; IPR002290; -;
 DR Pfam; PF00069; pkinase; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR SMART; SM00220; S_TKC; 1.
 KW ATP-binding; Calcium-binding; EGF-like domain; Glycoprotein; Repeat;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 595 AA; 63898 MW; 7B58FCA56E0641 CRC64;

Query Match 50.6%; Score 45; DB 10; Length 595;

Best Local Similarity 58.3%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12
 | ||||| |
 Db 20 GECNGRCGRVD 31

RESULT 7
 ID Q9LMP0 PRELIMINARY; PRT; 733 AA.
 AC Q9LMP0;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE F16F4.6 PROTEIN.
 GN F16F4.6.
 OS Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Liu S.X., Sakano H., Yu G., Etgu P., Lee J.M., Lenz C., Pham P.,
 RA Toriumi M., Byun E., Chan A., Chin C., Choi E., Chung M.,
 RA Goldsmith A., Gonzalez A., Liu A., Smith A., Vaysberg M., Altafi H.,
 RA Brooks S., Buehler E., Chao Q., Conn L., Conway A.B., Hansen N.F.,
 RA Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M., Nguyen M.,
 RA Palm C.J., Shinn P., Southwick A., Davis R.W., Ecker J.R.,
 RA Federspiel N.A., Theologis A.;
 RT "The sequence of BAC F16F4 from Arabidopsis thaliana chromosome 1.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AC036104; AAF81356.1; -;
 DR InterPro; IPR000152; -;
 DR InterPro; IPR000561; -;
 DR InterPro; IPR000719; -;
 DR InterPro; IPR001245; -;
 DR InterPro; IPR001881; -;
 DR InterPro; IPR002290; -;
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR SMART; SM00181; EGF; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase;
 KW Tyrosine-protein kinase.
 SQ SEQUENCE 733 AA; 81027 MW; DAA534D24C5A97AE CRC64;

Query Match 50.6%; Score 45; DB 10; Length 733;
 Best Local Similarity 58.3%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTSNGNGCCNYD 12
 | ||||| |
 Db 157 GECNGRCGRVD 168

RESULT 8
 ID O81820 PRELIMINARY; PRT; 735 AA.
 AC O81820;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE WALL-ASSOCIATED KINASE 1.
 GN WAK1.
 OS Arabidopsis thaliana (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA He Z.H., Cheeseman I., He D., Kohorn B.D.;
 RT "A cluster of five cell wall associated receptor kinase genes (Wak)
 are expressed in specific tissues of Arabidopsis.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AJ009696; CAA08794.1; -;
 DR Mendel; 32451; Arath;1197;32451.
 DR InterPro; IPR000152; -;
 DR InterPro; IPR000561; -;
 DR InterPro; IPR000719; -;
 DR InterPro; IPR001881; -;
 DR InterPro; IPR002290; -;
 DR Pfam; PF00069; pkinase; 1.
 DR PROSITE; PS001010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS0108; PROTEIN_KINASE_ST; 1.
 DR SMART; SM00220; STKc; 1.
 KW ATP-binding; Calcium-binding; EGF-like domain; Glycoprotein; Repeat;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 735 AA; 81210 MW; AAD41A28296093E6 CRC64;

Query Match 50.6%; Score 45; DB 10; Length 735;
 Best Local Similarity 58.3%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 GTSNGNGCCNYD 12
 Db 157 GECNGGCCRVDD 168

RESULT 9
 OI3304 PRELIMINARY; PRT; 790 AA.
 ID OI3304;
 AC OI3304;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE PROTEASE 1.
 OS Pneumocystis carinii f. sp. carinii.
 OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
 OC Pneumocystis.
 OX NCBI_TaxID=38081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97388577; PubMed=9245911;
 RA Lugli E.B., Allen A.G., Wakefield A.E.;
 RT "A Pneumocystis carinii multi-gene family with homology to subtilisin-
 like serine proteases.";
 RL Microbiology 143:0-0(0).
 DR EMBL; AF001304; AAB66701.1; -;
 DR HSP; P04189; ISCJ.
 DR MEROPS; S08.011; -;
 DR InterPro; IPR002884; -;
 DR InterPro; IPR002029; -;
 DR Pfam; PF01483; P; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR ProDom; PD000717; -; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 KW Protease.
 SQ SEQUENCE 790 AA; 88123 MW; C26F208D3647312D CRC64;

Query Match 50.6%; Score 45; DB 3; Length 790;

Best Local Similarity 60.0%; Pred. No. 18;
 Matches 9; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
 Qy 1 GTSNGNGC--CNYDG 13
 Db 349 GTGGGSLDGCNYDG 363
 RESULT 10
 Q9M9X1 PRELIMINARY; PRT; 883 AA.
 ID Q9M9X1;
 AC Q9M9X1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE PUTATIVE DISEASE RESISTANCE PROTEIN.
 GN F18C1.7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
 RA Renning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC F18C1 genomic sequence.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC011620; AAF26131.1; -;
 DR InterPro; IPR001611; -;
 DR InterPro; IPR003592; -;
 DR Pfam; PF00560; LRR; 15.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00370; LRR; 1.
 SQ SEQUENCE 883 AA; 98503 MW; EA32DFE61B2E2525 CRC64;

Query Match 50.6%; Score 45; DB 10; Length 883;
 Best Local Similarity 60.0%; Pred. No. 20;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 4 NGNGCCNYDG 13
 Db 65 NSDCCCHWDG 74

RESULT 11
 Q9W2B6 PRELIMINARY; PRT; 109 AA.
 ID Q9W2B6;
 AC Q9W2B6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE CG13494 PROTEIN.
 GN CG13494.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RA Science 287:2185-2195(2000).
 DR EMBL: AE003455; AAF46776.1; -.
 DR FlyBase: FBgn0034671; CG13494.
 DR FlyBase: FBgn0034671; CG13494.
 SQ SEQUENCE 109 AA; 10413 MW; 84E0A6C087990895 CRC64;

Query Match 49.4%; Score 44; DB 5; Length 109;
 Best Local Similarity 53.8%; Pred. No. 4.2;
 Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GTSNGNGCCNDVG 13
 I I I I I I I
 Db 72 GEGGGGGCGGDDG 84

RESULT 12
 Q9DFE9
 ID Q9DFE9 PRELIMINARY; PRT; 165 AA.
 AC Q9DFE9;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 18.7 KDA PROTEIN (FRAGMENT).
 OS *Oncorhynchus mykiss* (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bayne C.J., Gerwick L., Fujiki K., Nakao M., Yano T.;
 RA "Immune-relevant (including acute phase) genes identified in the
 RT livers of rainbow trout, *Oncorhynchus mykiss*, by means of suppression
 RT subtractive hybridization.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF281341; AAG30016.1; -.
 KW Hypothetical protein.
 FT NON_TER 1
 FT NON_TER 165
 FT NON_TER 165
 SQ SEQUENCE 165 AA; 18682 MW; A93EB87C38CF4502 CRC64;

Query Match 49.4%; Score 44; DB 13; Length 165;

Best Local Similarity 66.7%; Pred. No. 6.1;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 GTSNGNGCCNDVG 12
 I I I I I I I I
 Db 77 GTRGGNGRCNCD 88
 RESULT 13
 Q38326
 ID Q38326 PRELIMINARY; PRT; 258 AA.
 AC Q38326;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 28.7 KDA PROTEIN (ORF258).
 OS *Lactococcus phage BK3-T*.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 OX NCBI_TaxID=31754;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Boyce J.D., Davidson B.E., Hillier A.J.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L44593; AAA98586.1; -.
 SQ SEQUENCE 258 AA; 28714 MW; 37AD2875724D59D4 CRC64;

Query Match 48.3%; Score 43; DB 9; Length 258;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GTSNGNGCCNDVG 12
 I I I I I I I I
 Db 183 GDSNGNGGTNNY 194

RESULT 14
 Q9FN37
 ID Q9FN37 PRELIMINARY; PRT; 1036 AA.
 AC Q9FN37;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE RECEPTOR PROTEIN KINASE-LIKE PROTEIN.
 OS *Arabidopsis thaliana* (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=98162728; PubMed=9501997;
 RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RA "Structural analysis of *Arabidopsis thaliana* chromosome 5. III.
 RT Sequence features of the regions of 1,191,918 bp covered by seventeen
 RT physically assigned pl. clones.";
 RL DNA Res. 4:401-414(1997).
 DR EMBL: AB007644; BAB10719.1; -.
 KW Kinase.
 SQ SEQUENCE 1036 AA; 114339 MW; 6DF9511FC2A4E261 CRC64;

Query Match 48.3%; Score 43; DB 10; Length 1036;
 Best Local Similarity 60.0%; Pred. No. 51;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 NGNGCCNDVG 13
 I I I I I I I I
 Db 45 NGSRCCWDG 54

```

RESULT 15
Q9M8N2 PRELIMINARY; PRT; 198 AA.
AC Q9M8N2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE T21F11.2 PROTEIN.
GN T21F11.2
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome I BAC T21F11 genomic sequence.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC018849; AAF27126.1; -.
DR InterPro; IPR001159; -.
DR Pfam; PF00035; dsrm; 1.
DR PROSITE; PS0137; DS_RBD; 1.
DR SMART; SM00358; DSRM; 1.
SQ SEQUENCE 198 AA; 22619 MW; 5AD2F1799C4A0791 CRC64;

```

```

Query Match 47.2%; Score 42; DB 10; Length 198;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 8 CCNYDGP 14
DB 135 CCNVDGP 141

```

Search completed: November 21, 2001, 16:27:40
Job time: 404 sec